

DAS BONNER ZEITUNGSKORPUS TEIL I

INFORMATIONEN FÜR DEN BENUTZER

0. Vorbemerkung
1. Zielsetzung und Textauswahl
2. Aufbau und Zusammensetzung
3. Zur Codierung des BZK 1
4. Zur Identifikation und Klassifikation der Zeitungsartikel
5. Möglichkeiten der Auswertung
6. Ausblick
7. Literatur

0. VORBEMERKUNG

Über das Bonner Zeitungskorpus ist, wie die Literaturübersicht am Schluß zeigt, an verschiedenen Stellen berichtet worden, am ausführlichsten in dem Sammelband Hellmann (Hg.) 1984, vor allem dort im Bericht II von B. Schaefer. Allerdings bezieht sich jener Bericht auf die Bonner Version des BZK. Inzwischen ist das BZK 1 vollständig überarbeitet worden. Es gab noch zwei weitere Gründe, eine neue - und zwar vollständige - Beschreibung des BZK vorzulegen: Zum einen ist der erwähnte Sammelband so exorbitant teuer, daß er nicht überall erreichbar sein wird, zum anderen ist mit erhöhtem Benutzerinteresse zu rechnen, da das BZK Teil 1 jetzt gleichzeitig im MCS-Verlag auf Microfiches erscheint und im IdS mit dem Abfragesystem REFER auswertbar ist.

Ich danke den Kolleginnen und Kollegen im IdS, die sich in den letzten 3 Jahren mit mir der Mühe unterzogen haben, ein Drei-Millionen-Korpus zu überarbeiten und bereitzustellen, insbesondere Willi Oksas, Snjezana Benaković und Peter Mückenmüller.

1. ZIELSETZUNG UND TEXTAUSWAHL

1.1. Gegenstandsbestimmung

Das Bonner Zeitungskorpus (BZK) ist eine systematisch zusammengestellte, maschinenverfügbar gemachte Sammlung von Texten aus bestimmten Tageszeitungen der Bundesrepublik Deutschland und der Deutschen Demokratischen Republik, die zum Zwecke linguistischer Untersuchungen zur deutschen Gegenwartssprache, insbesondere zum Zweck vergleichender Untersuchungen des öffentlichen Sprachgebrauchs in den beiden deutschen Staaten, angelegt wurde. Das BZK Teil 1 ist eine bestimmte Teilmenge davon (nämlich aus der WELT und dem NEUEN DEUTSCHLAND), und zwar diejenige, die zur Zeit überarbeitet, an die Textkonventionen des IdS Mannheim angepaßt und mittels vorhandener Auswertungsprogramme auswertbar ist (vgl. 2.3.).

1.2. Anlaß und Zielsetzung

Das IdS hatte 1964 in Bonn eine Forschungsstelle eingerichtet, deren Hauptaufgabe die Untersuchung der sprachlichen Differenzierungen zwischen den beiden deutschen Staaten sein sollte. Die damalige Forschungssituation legte es nahe, sich zunächst das Ziel zu setzen, eine ausreichende und von subjektiven Auswahlkriterien möglichst unabhängige Grundlage für empirische Untersuchungen zu schaffen, und zwar nach Lage der Dinge vor allem für Untersuchungen zum Wortschatz und Wortgebrauch.

Eine solche Textgrundlage sollte auf jeden Fall maschinell verfügbar gemacht werden, und zwar einerseits zur Entlastung von zeitraubenden Auswertungs-, Sortier- und Auflistungsarbeiten, zum anderen aber um Häufigkeit und Verteilung sprachlicher Phänomene als Erkenntnis- und Beschreibungskategorie nutzen zu können. Darüber hinaus sollte das zu schaffende Textkorpus folgende Eigenschaften aufweisen: Es sollte

- aktuell sein
- wenig autorenspezifisch sein
- thematisch weit gestreut sein
- dem Forschungsgegenstand angemessen, also möglichst ergiebig für Ost-West-Unterschiede sein.

Da für uns die Konstituierung eines großen Mischkorpus aus unterschiedlichen Publikationsgattungen aus Kosten- und organisatorischen Gründen nicht in Frage kam, mußten wir uns also für eine bestimmte Publikationsgattung entscheiden. Am ehesten werden die gestellten Anforderungen von Tageszeitungen erfüllt, und zwar von solchen, die sich an gehobene Ansprüche richten (also nicht von Boulevardblättern).

1.3. Auswahl und Zusammensetzung

Textauswahl und Aufnahmeprinzip

Die Kriterien für die Auswahl eines Textkorpus werden - bewußt oder unbewußt - bestimmt durch den Zweck, für den das Korpus aufgenommen werden soll. Dies gilt schon für die Wahl der Grundgesamtheit: Wenn ich öffentlichen Sprachgebrauch, und zwar in den beiden deutschen Staaten, untersuchen will, ist die Wahl von überregionalen Tageszeitungen, und zwar aus Ost und West, vielleicht nicht die einzig mögliche, aber sicher eine zweckmäßige Grundlage für den Aufbau eines Korpus.

Die nächste Frage gilt dem Aufnahmeprinzip: Zugespitzt (es gibt Übergangsformen) lautet die Alternative: Benötige ich ein Korpus für eine einzige, vorher bekannte Fragestellung oder ein Korpus für eine Vielzahl verschiedener, z.T. vielleicht noch nicht bekannter Fragestellungen?

Diesen sehr unterschiedlichen Zwecken entsprechen auch - auch hier wieder zugespitzt - sehr unterschiedliche Aufnahmeprinzipien:

- (a) thematisch gesteuerte Auswahl
- (b) statistisch gesteuerte Auswahl

Bei ersterem Aufnahmeprinzip wähle ich aus der festgelegten Grundgesamtheit solche Texte aus, die - nach meinem Vorverständnis vom Forschungsgegenstand und vom Thema - besonders ergiebig und interessant zu sein scheinen. Dieses Auswahlverfahren ist sehr effektiv, denn ich komme mit relativ geringem Aufwand zu einer Auswahl, die in hoher Dichte die von mir zu untersuchenden Erscheinungen enthält und daher für diesen Gegenstand sehr gut geeignet ist, - für alle anderen Fragestellungen jedoch

voraussichtlich entsprechend schlecht. Hinzuweisen ist noch auf zwei weitere Punkte: Eine rein thematisch gesteuerte Auswahl erlaubt keine statistisch abgesicherten Rückschlüsse auf die Verhältnisse in der Grundgesamtheit, der die Auswahl entnommen ist, und sie erlaubt keinerlei Vergleiche mit anderen thematisch gesteuerten Auswahlen. Zweitens: Es besteht die Gefahr des methodischen Zirkels: Ich erhalte im wesentlichen das bestätigt, was ich über den Gegenstand kategorial schon vorher wußte.

Das zweite Aufnahmeprinzip, das ja von der Vorstellung einer Vielzahl verschiedener Benutzer mit sehr unterschiedlichen Fragestellungen ausgeht, erlaubt von vornherein keine Vorweg-Auswahl des "Wichtigeren" oder "Interessanteren", da jede qualitative Auswahl eine Einschränkung der Benutzung durch andere bedeuten kann. Konsequenterweise wird man versuchen, die Auswahl so zu treffen, daß sie alle bekannten Eigenschaften der Grundgesamtheit möglichst getreu widerspiegelt. Beachtet man bei der Zusammenstellung der Auswahl gewisse systematische Grundsätze, so kann die Auswahl als in jeweils bestimmbar Grenzen repräsentativ in bezug auf die Grundgesamtheit angesehen werden, d.h. Zahlenangaben zu Vorkommen oder Verteilung bestimmter Erscheinungen in der Auswahl lassen mit hoher Wahrscheinlichkeit den Schluß auf entsprechende Verhältnisse in der Grundgesamtheit zu. Außerdem sind alle Teile des Korpus, die nach dem gleichen Verfahren ausgewählt und aufgenommen worden sind, auch untereinander vergleichbar.

Der Nachteil bei diesem Auswahlverfahren ist, daß man, um auch für speziellere Untersuchungen noch genügend Material zur Verfügung stellen zu können, sehr viel mehr Text als bei einer thematisch gesteuerten Auswahl aufnehmen muß. Außerdem verlangt eine so breit gestreute Auswahl einen wesentlich höheren Aufwand an Klassifikation der einzelnen Texteinheiten, um dem künftigen Benutzer eine Selektion der Texte im Hinblick auf bestimmte gewünschte oder nicht gewünschte Merkmale zu ermöglichen.

Wir entschieden uns im Jahre 1964/65 für ein rein statistisch gesteuertes Aufnahmeprinzip, und zwar aus folgenden Gründen:

Erforschung des öffentlichen Sprachgebrauchs auf der Basis von Tageszei-
tungen war damals und ist auch heute keineswegs nur unter Ost-West-
Aspekt interessant. Es war vorauszusehen, daß die Linguistik, wäre ein
geeignetes Korpus erst einmal vorhanden, dies ist auch unter generellen
Aspekten des Wortschatzes, des Wortgebrauchs, des Stils, der Wortbil-
dung, der Syntax usw. auswerten würde. Angesichts des Aufwandes an Zeit
und Kosten, den jede maschinelle Dokumentation von Texten - gleich ob
thematisch gesteuert oder statistisch gesteuert - verursacht, schien es
uns auch ökonomischer, die Textauswahl nicht auf den einen Forschungsge-
genstand hin zu beschneiden. Im übrigen sahen wir in der Tatsache, daß
das IdS eine zentrale Forschungsinstitution ist, eine gewisse Verpflich-
tung, ein neu aufzubauendes Korpus nicht als ad-hoc-Korpus für über-
wiegend hausinterne Zwecke anzulegen, sondern es für eine möglichst
breite wissenschaftliche Nutzung innerhalb u n d außerhalb des IdS zu
konzipieren. Denn der Aufbau eines ad-hoc-Korpus für eine spezielle Fra-
gestellung führt oft zum Aufbau anderer ad-hoc-Korpora für ähnliche,
ebenfalls spezielle Fragestellungen an anderer Stelle, ohne daß sich
durch die Kumulation von ad-hoc-Korpora ein in sich vergleichbares
Korpus deutscher Gegenwartssprache - oder auch nur eines Ausschnitts
daraus - ergäbe.

Über einen Nachteil waren wir uns seinerzeit allerdings im klaren: Eine
Korpuskonzeption, die im skizzierten Sinne auf breite verschiedenartige
Nutzung und statistische Repräsentativität zur jeweiligen Grundgesamt-
heit hin angelegt ist, kann zwar für die meisten Fragestellungen ge-
eignetes, aber für manche davon sicherlich nicht genügend Material be-
reitstellen, - jedenfalls wenn man von einigermaßen realistischen Kos-
ten- und Zeitvorstellungen ausgeht. Die Konzeption mußte also so be-
schaffen sein, daß sie Materialerweiterungen und -verdichtungen in ver-
schiedener Hinsicht jederzeit zuläßt.

Zum Problem der Aktualität

Jedes Korpus beginnt vom Tage seiner Verfügbarkeit an zu veralten. Nun
ist der Begriff "Deutsche Gegenwartssprache" zwar recht dehnbar - man
setzt den Beginn meist 1945 oder 1949 und läßt ihn jeweils in der Gegen-

wart enden -, jedoch hat sich für eine Reihe von Untersuchungen insbesondere zum Wortschatz und zum Stil ein engerer Begriff von "Gegenwart" und zugleich ein Bedürfnis nach stärkerer Differenzierung der Entwicklung in der Nachkriegszeit herausgestellt. Die Bonner Korpuskonzeption wollte dem Rechnung tragen.

Man kann Aktualität dadurch erreichen, daß man laufend Texte aufnimmt. Dies ist bei Periodika, die ohnehin parallel zum Zeitkontinuum permanent Texte produzieren, grundsätzlich kein Problem. Allerdings stößt man sehr schnell auf Kapazitätsgrenzen. Westdeutsche überregionale Tageszeitungen produzieren pro Jahrgang etwa 5.000 Seiten "Normal"-Text, dazu nochmals zwischen 1.500 und 3.000 Seiten Beilagentext. Bei rd. 3.000 lfd. Wörtern pro Seite ergibt sich daraus eine Menge von 15 Mio. lfd. Textwörtern zusätzlich ca. 6 Mio. lfd. Wörtern aus Beilagen. Solche Mengen führen schon bei e i n e r Zeitung und e i n e m Jahrgang ins Absurde, um so mehr bei einer beabsichtigten Aktualisierung durch laufende Aufnahme von Texten. Ein Ausweg ist die Aktualisierung durch Aufnahme von Querschnitten in regelmäßigen Intervallen. Solange dieses Verfahren weitergeführt wird, ist Aktualität in hohem Maße gewährleistet; gleichzeitig sind die Voraussetzungen gegeben für diachronische Untersuchungen der Sprachentwicklung mit zunehmender zeitlicher Tiefe.

Die Entscheidung für die Textaufnahme in Form diachronischer Querschnitte in bestimmten zeitlichen Intervallen nimmt natürlich Lücken in Kauf. Sachlich und sprachlich können dabei wichtige Erscheinungen undokumentiert bleiben. Unsere Entscheidung für ein fünfjähriges Aufnahmeintervall läßt z.B. Ereignisse wie den 17. Juni 1953, den 13. August 1961, die Verhandlungen über das Berlin-Abkommen 1970-72, den Wahlsieg der sozial-liberalen-Koalition oder auch die Olympiade in München unberücksichtigt; andere wie z.B. die Gründung der beiden deutschen Staaten, die Bildung der großen Koalition, die Guillaume-Krise und der Wechsel Brandt/Schmidt, die Olympiade von Tokio (1964) u.v.a. sind im Bonner Zeitungskorpus sehr gut vertreten.

Es spricht also - vom Kosten- und Kapazitätsargument abgesehen - einiges für kürzere Intervalle. Es spricht noch mehr gegen längere Intervalle.

98

Wir halten Fünf-Jahres-Intervalle für gerade noch ausreichend: mittelfristige Erscheinungen haben noch eine ausreichende Wahrscheinlichkeit, in der Auswahl dokumentiert zu werden; kürzerfristige nur dann, wenn sie in einen der dokumentierten Jahrgänge oder doch in dessen Nähe fallen.

Im konkreten Fall haben wir mit dem Jahrgang 1964 deshalb begonnen, weil er bei Beginn der Textdokumentation (im Frühjahr 1965) der erste und aktuellste abgeschlossen vorliegende Jahrgang war; damit war auch gewährleistet, daß unsere Auswahl im Jahr der Gründung der beiden deutschen Staaten, 1949, beginnen würde.

Zur Auswahl innerhalb des Jahrgangs

Wie schon festgestellt, ist die vollständige Aufnahme eines Jahrgangs aus Kapazitätsgründen unmöglich. Eine Auswahl aus dem Jahrgang kann in verschiedener Hinsicht erfolgen:

Man kann bei gleichem Aufwand z. B. sowohl einen Monat vollständig als auch über den ganzen Jahrgang hin jede zwölfte Ausgabe erfassen. Für bestimmte Untersuchungen, z.B. eine Untersuchung zur Sprache der Fußball-Sportberichterstattung, kann die erstere Variante eine hervorragende Materialbasis mit sehr hoher Dichte ergeben, sofern man darauf achtet, daß der ausgewählte Monat nicht gerade in die Zeit der Fußballferien fällt. Unter der Voraussetzung sehr unterschiedlicher Fragestellungen macht die Entscheidung für einen bestimmten Erfassungszeitraum jedoch große Begründungsschwierigkeiten. Denn jede Untersuchung von Zeitungen bestätigt, daß Zeitungen im Verlauf eines Jahres sehr starke, regelmäßige "thematische Dominanzen" aufweisen, so insbesondere im Bereich des Sports, der Wirtschaft (vor allem Landwirtschaft), der Sparte Reise/Erholung, der Werbung, aber auch in Teilen der Sparte Kultur (Theaterferien) und teilweise auch Politik (thematische Schwerpunkte z.B. am Jahresanfang bzw. -ende, zu bestimmten Gedenktagen etc.). Ein Textkorpus, das Zeitungssprache widerspiegeln will, muß diese thematischen Dominanzen angemessen berücksichtigen, d.h. es muß sich auf den ganzen Jahrgang einer Zeitung beziehen.

Die Stichproben, die aus dem Kontinuum eines Jahrgangs entnommen werden sollen, sollten in gleichmäßigen, nicht zu großen Abständen über den ganzen Jahrgang verteilt werden.

Uns schien die Berücksichtigung von je einer Ausgabe pro Woche dafür am geeignetsten.

Außer im Ablauf eines Jahres gibt es auch im Ablauf einer Woche bei Tageszeitungen sehr deutliche thematische Dominanzen. So gibt es einen ausgedehnten Sportteil montags, ausführliche Kinoprogramme dienstags oder freitags, Wohnungsmarkt vor allem mittwochs und samstags; bestimmte Beilagen erscheinen an bestimmten Wochentagen; vor allem die Samstagsausgabe weist eine Reihe von Besonderheiten auf. Es war also notwendig, durch einen entsprechenden Turnus der Aufnahme dafür zu sorgen, daß alle Wochentagsausgaben in der Jahrgangsauswahl berücksichtigt wurden.

Dies führte zu folgendem rotierenden Turnus: Zum Beispiel erste Jahreswoche: Montagsausgabe, zweite Woche Dienstagsausgabe, dritte Woche Mittwochsausgabe ..., fünfte Woche Freitagsausgabe, sechste Woche Samstagsausgabe, siebente Woche Montagsausgabe.

Während der Normalabstand zwischen zwei Ausgaben also acht Tage beträgt, beträgt er zwischen der Samstags- und Montagsausgabe nur zwei Tage, so daß ein durchschnittlicher Abstand von knapp sieben Tagen gewährleistet ist. Insgesamt führt dies zu einer Berücksichtigung von 52 oder 53 Ausgaben pro Jahr.

Zur Auswahl der Seiten

Falls man diese 53 Ausgaben vollständig aufnehmen kann, entfallen weitere Auswahlüberlegungen. Für uns lag die damit anfallende Menge von ca. 700-800 Seiten pro Zeitung und Jahrgang jedoch noch weit jenseits unserer Möglichkeiten. Eine Seitenauswahl war also unumgänglich. Nun gibt es auch innerhalb einer jeden Ausgabe Regelmäßigkeiten, die zu beachten sind, wenn man bestimmte Seiten auswählen will. Während die erste und die letzte Seite bei allen Zeitungen in aller Regel eine Mischung ver-
100

schiedener Themen und Sachgebiete aufweisen, sind die übrigen Seiten meist in sog. "Sparten" thematisch gebündelt. Diese Bündelung ist zwar nicht starr, weist jedoch jeweils zeitungstypische Regelmäßigkeiten auf. Diese Regelmäßigkeiten müssen sich in der Auswahl widerspiegeln, d. h. jede Seitenklasse soll in der Auswahl anteilig repräsentiert sein.

Zur Aufnahmequote

Wenn man davon ausgeht, daß aus jeder Woche eine Ausgabe und aus jeder Ausgabe nur eine Seite aufgenommen wird, ergibt dies eine Mindestauswahlmenge von 52 oder 53 Seiten. Mit diesem Minimum haben wir in der Tat bei den Jahrgängen ND 64 und ND 54 begonnen (die Auswahl ND wurde später durch Beilagenseiten erweitert). Bei einer Grundgesamtheit von 2125 Seiten im Jahrgang ND 64 und 1976 Seiten bei ND 54 ergaben sich somit Aufnahmequoten von 2,5 % (ND 64) bzw. 2,6 % für ND 54.

Wir haben im folgenden eine Auswahlquote von ca. 2,4 % bis 2,5 % für alle Jahrgänge des BZK I konstant gehalten, mindestens aber 52 Seiten pro Jahrgang aufgenommen (zur Ausnahme WE 49 siehe unten 2.3., Anm. 1).

Da die WELT weit mehr Seiten und Text produziert als das ND, ergibt sich für die WELT bei gleicher Quote eine größere Auswahlmenge und zugleich eine höhere Aufnahmedichte pro Ausgabe.

Das Schema sieht folgendermaßen aus:

z.B. ND 74

1. Woche Mi	02.01.	S. 1
2. Woche Do	10.01.	S. 2
3. Woche Fr	18.01.	S. 3
4. Woche Sa	26.01.	S. 4
5. Woche Mo	28.01.	S. 5
6. Woche Di	05.02.	S. 6, 3
usw.		

z.B. WELT 74

Mi	S. 1, 2, 3
Do	S. 2, 3, 4
Fr	S. 3, 4, 5
Sa	S. 32, 33, 34
Mo	S. 5, 6, 7
Di	S. 6, 7, 8
usw.	

Zur Bestimmung der Grundgesamtheit

Nicht alles, was eine Zeitung im Laufe eines Jahres produziert bzw. an ihre Leser verteilt, ist in gleicher Weise Bestandteil dieser Zeitung. Wir unterscheiden

- Grundmenge 1: Redaktioneller Teil der Zeitung und regelmäßige Beilagen, soweit sie normal durchlaufend numeriert sind.
- Grundmenge 2: gesondert (meist römisch) numerierte Beilagen, die regelmäßiger Bestandteil bestimmter Wochentagsausgaben dieser Zeitung sind (z.B. Stellenanzeigen in der Samstagsausgabe, die "Geistige Welt", die "Gebildete Nation" (ND)),
- Grundmenge 3: nach Druckbild, Numerierung und Aufmachung eigenständige Beilagen ("Sonderbeilagen"), die nicht regelmäßiger Bestandteil dieser Zeitung sind oder als "Zeitung in der Zeitung" erscheinen (z.B. bei der WELT die "Welt der Literatur", die Rundfunk-Fernseh-Beilage etc.).

Zur besseren Unterscheidung der Grundmengen 1 und 2 wird folgende Zusatzbestimmung verwendet: Wenn eine Beilage regelmäßig mehr als die Hälfte des redaktionellen Teils ausmacht, wird sie zur Grundmenge 2 gerechnet.

Auf der Grundlage der genauen Auszählung aller Ausgaben des Jahrgangs wird ermittelt, welche Seiten wie oft in den Grundmengen vorkommen. Diese Grundmengen werden in unterschiedlichen Quoten repräsentiert, und zwar wird

- die Grundmenge 1 mit einer Quote von ca. 2,4 %
- die Grundmenge 2 mit einer Quote von ca. 1,5 bis 1,8 %
- die Grundmenge 3 überhaupt nicht repräsentiert.

Als Grundgesamtheit ist damit die Summe aller Zeitungsseiten (=Elemente) der Grundmengen 1 und 2 eines Jahrgangs definiert.

94. AUFBAU UND ZUSAMMENSETZUNG

2.1 Aufbau des Bonner Zeitungskorpus

Mengenangaben (in lfd. Wörtern) gerundet

Doppelt umrandete Felder: = BZK Teil 1

		WE 49 Ausgabe Nord (163.000)	ND 49 Ausgabe R (206.000)		
		WE 54 Ausgabe F (258.000)	ND 54 Ausgabe B (208.000)		
		WE 59 Ausgabe D (293.000)	ND 59 Ausgabe B (208.000)		
	BG 64 (340.000)	WE 64 Ausgabe B (379.000)	ND 64 Ausgabe B (194.000)	NN 64 (100.000)	
		WE 69 Ausgabe D (414.000)	ND 69 Ausgabe B (174.000)		
FR 74 (310.000)	BG 74 (340.000)	WE 74 Ausgabe D (363.000)	ND 74 Ausgabe R (179.000)	NN 74 (100.000)	MO 74 (140.000)
310.000	680.000	3,039.000	1,169.000	200.000	140.000

BRD-Zeitungen

WE = DIE WELT (B = Berliner Ausgabe, D = Deutschland-Ausgabe (Essen), F = Norddeutsche Ausgabe (Hamburg))

GB = Bonner-General-Anzeiger (Stadttausgabe)

FR = Frankfurter Rundschau

DDR-Zeitungen

ND = Neues Deutschland (B = Berliner Ausgabe, R = Republik-Ausgabe)

NN = Norddeutsche Neueste Nachrichten (NDPD, Rostock)

MO = Der Morgen (LDPD, Berlin Ost)

Umfang des Bonner Zeitungskorpus Teil 1: 3,04 Mio. lfd. Wörter

Teil 2: 1,33 Mio. lfd. Wörter

2.2. Versionen des BZK

Das Bonner Zeitungskorpus wurde in 3 Phasen geschaffen:

(1) 1965-75 (mit Unterbrechungen) Jg. WE 54, 64, 59
Jg. ND 54, 64, 69
und teilweise ND 74

(2) 1976-78 (Projekt "Ost-West-Wortschatz") Jg. WE 49, 59, 74
Jg. ND 49, 59, 74

ferner	Jg. BG	64, 74
	Jg. FR	74
	Jg. NN	64, 74
	Jg. MO	74

Die Texte der Phase (2) wurden mit Unterstützung der Deutschen Forschungsgemeinschaft extern erfaßt und einmal Korrektur gelesen.

(3) 1981-84

Überarbeitung des Teils 1 (WE und ND) und Anpassung an die Mannheimer Textkonventionen

Jg. WE 49, 54, 59, 64, 69, 74

Jg. WE 49, 54, 59, 64, 69, 74

Die Texte der Phase 1 und 2 - "Bonner Version" - weisen gegenüber der Phase 3 - "Mannheimer Version" - einige Besonderheiten auf:

1. Sie sind weniger gut durchkorrigiert.

2. Sie weisen Schwankungen in der Codierung auf (bedingt u.a. durch den Übergang zur externen Texterfassung in Phase 2).
3. Sie sind komplizierter transkribiert (vgl. Punkt 3.2.).
4. Umlaute und B sind als **ae, oe, ue, sz** codiert.
5. Sie haben eine andere Satzschlußcodierung (vgl. Punkt 3.2.).
6. Die Texte der Phase 1 sind mit Artikelnummer/Zeilennummer nummeriert; die Texte der Phase 2 sind durchlaufend zeilenweise nummeriert.
7. Insbesondere die Texte der Phase 1 weisen eine erweiterte Klassifikation auf (vgl. Punkt 4.1).

Für bestimmte Fragestellungen eignen sich die BZK-Texte in der Bonner Version (trotz ihrer höheren Fehlerquote) besser als die Mannheimer Version, daher bleibt auch die Bonner Version auf Magnetband gespeichert. Solche Fragestellungen können z.B. sein: Untersuchungen zu Personennamen samt Titulaturen, zu Sach-Eigennamen, zur Markierung von Satzschlüssen in Zeitungstexten.

Die Mannheimer Version steht zur Verfügung

- in satznumerierter Form zur Auswertung mittels REFER (vgl. Punkt 5.2.3)
- in zeilennumerierter Version auf Magnetband
- zusätzlich auf Mikrofiches (vgl. Punkt 5.2.2).

Auch die Mannheimer Version ist nicht absolut fehlerfrei; dies wäre bei einem manuell erfaßten Drei-Millionen-Korpus nur mit unvertretbar hohem Aufwand erreichbar. Wir haben eine Fehlerquote von durchschnittlich 1 Fehler auf 3 Seiten Text (à 60 Zeilen) in Kauf genommen. Für ND und WELT 1949 sowie für WELT 1969 wurde dies Ziel wahrscheinlich nicht ganz erreicht (1 Fehler auf 2 Seiten Text).

2.3. Umfang der Grundmenge 1 und Seitenauswahl

Bonner Zeitungskorpus Teil 1

Jg.	Grundmenge 1	Seitenauswahl	Quote
WE 49	1.684	41 (54)	2,43 (3,2) ⁽¹⁾
WE 54	3.348	80	2,40
WE 59	3.920	95	2,42
WE 64	5.635	135	2,40
WE 69	6.988	168	2,40
WE 74	6.243	150	2,40
Gesamt			
WELT	27.818	669	2,40
ND 49	1.814	52	2,87
ND 54	1.976	52	2,63
ND 59	2.300	56	2,43
ND 64	2.125 (+312)	53 (+8)	2,50 ⁽²⁾
ND 69	2.504	61	2,44
ND 74	2.484	61	2,46
Gesamt			
ND	13.515	343	2,54
WE + ND	41.333	1012	2,45

Die 1012 Seiten Auswahlmenge des BZK 1 repräsentieren somit eine Grundgesamtheit von 41.333 Seiten in einer Quote von durchschnittlich 2,45 %.

Anm. (1): Aus der turnusmäßig ermittelten Auswahlmenge von 54 Seiten wurden 13 herausgenommen und als Zusatzmenge gespeichert, da die WELT im 1. Halbjahr 49 nur an jedem 2. Tag erschienen ist.

Anm. (2): Die Beilage "Die Gebildete Nation" wurde nachträglich in die Grundmenge und die Auswahl einbezogen.

Zusätzlich zu diesen 1012 Seiten wurden zu den meisten Jahrgängen weitere Seiten als Zusatzmengen aufgenommen. Ihr Vorhandensein hat verschiedene Gründe. Bei den früheren Bonner Jahrgängen der Phase 1 handelt es sich größtenteils um eine zusätzliche Auswahl aus dem Monat Februar (sog. "Februarserie"), die aufgenommen wurde, um bestimmte Fragen statistischer Repräsentativität bei erhöhter Aufnahmequote zu untersuchen. Teilweise handelt es sich um irrtümlich und später ausgetauschte Seiten. In der Mehrzahl jedoch, insbesondere bei den Texten der Phase 2, handelt es sich um ausgewählte Seiten aus der Grundmenge 2, also um Beilagen-Seiten. Die jeweiligen Zusatzmengen sind in 2.4.1 und 2.4.2 aufgeschlüsselt.

Die Texte dieser Zusatzmengen liegen zum größeren Teil nur in der Bonner Version vor, d.h. sie sind nicht an die Mannheimer Codierungs-Konvention angepaßt, nur einmal korrigiert und sind auch nicht mittels REFER am Bildschirm, sondern nur über allgemeine Suchwortprogramme auswertbar.

2.4. Aufschlüsselung der ausgewählten Seiten nach dem Inhalt

Vorbemerkung

Die Aufschlüsselung muß für ND und WELT gesondert vorgenommen werden, da der Spartenaufbau bzw. die Ressortgliederung stark differieren. Aber auch innerhalb derselben Zeitung schwankt diese Gliederung stark. Die Zuordnung der Seiten zu einem einmal festgelegten Schema ist daher zumindest bei WE 49 und 54 oft schwierig.

Die Zählung berücksichtigt bei Seiten mit gemischtem Inhalt nur jeweils halbe Seiten (z.B. 1/2 Wirtschaft, 1/2 Roman).

Weitergehende Mischungen (z.B. zusätzlich noch Anzeigen) wurden nicht berücksichtigt. Generell können Anzeigen auf jeder Seite vorkommen; gezählt wurden sie nur, wenn sie annähernd die Hälfte oder mehr einer Seite ausmachen.

2.4.1. Spartenverteilung DIE WELT

Sparte	WE 49	WE 54	WE 59	WE 64	WE 69	WE 74
Titelseite	6	8	8	8	8	7
Politik	6	11	23	27	26	23 1/2
Meinung					6	7
Reportage		3		1	6 1/2	5 1/2
Anzeigen	4 1/2	13 1/2	18 1/2	24	34 1/2	44
Leserbriefe				4	4	4
Sport	3	4 1/2	6	8	14	8
Wirtschaft	5	13	16 1/2	27	25	17
Börse		1	4 1/2	5	13 1/2	12
Kultur/Feuilleton	8	15	5	5	8	16 1/2
Roman		2	3 1/2	5	6 1/2	
Wissenschaft	1					3 1/2
Motor/Reise			1	3		
Aus aller Welt		9	7	3	8	2
Berlin/Regionales			2	15		
Nachrichten/Berichte (s.Anm.)	7 1/2					
Summe	41	80	95	135	168	150
Zusatzauswahl						
Titelseite	2			2		
"Februarmenge" (gem.)		4		12		
Kultur/Feuilleton	2 1/2		4 1/2	1 1/2		
Anzeigen/Stellen-				8		
Anzeigen sonst.	2		2 1/2			
Wirtschaft	2 1/2					
Beilage GW			5	4		5
Beilage Humor			2	2		1
Beilage Frau			2	1		
Sonstiges	2	1	2			
Reise/Erholung				3 1/2		1
Politik	2			3		
Summe	13	5	18	37	0	7

Anmerkung zu WE 49:

In WE 49 ist eine Sparten- oder Ressortenteilung kaum erkennbar. In der Mehrzahl der Seiten werden Themen aus Innen- und Außenpolitik, Wirtschaft und Gesellschaft bunt gemischt. Sie werden für diesen Sonderfall in der Rubrik Nachrichten/Berichte (7 1/2) zusammengefaßt.

2.4.2 Spartenverteilung NEUES DEUTSCHLAND

Sparte	ND 49	ND 54	ND 59	ND 64	ND 69	ND 74
Titelseite	9	8	9	7	7	8
Politik	11	18	14 1/2	14	17	17 1/2
Wirtschaft/Produktion	5	3 1/2	5 1/2	3		3 1/2
Kultur/Feuilleton/ Wissenschaft	9	3	7	4+5	5	6 1/2
Reden/Reportagen	6 1/2	11	10	11	13 1/2	12 1/2
Leserbr./Diskuss.				1+2	1	1
Sport	3 1/2	3	4 1/2	8	4	6
Berlin/Lokales	5	4	3 1/2		4 1/2	4 1/2
Anzeigen	1	1	3	2	6	1 1/2
Vermischtes ²⁾	2			3+1		
Summe	52	52	56	53+8	61	61
Zusatzauswahl					Z E	
Reden				18	1 2	
"Februarserie"(div.)		5		14		
Kultur		1	3		2 1	2
Wissenschaft/Technik			1		2	2
Reportagen/Reden					1	2
Beilage Freizeit			1		1	
Beilage Lokales					1	
Gemischte Auswahl		8		1	1 4	1
Anzeigen		1				1
Summe	0	15	5	33	8 8	8

Anmerkung:

ND 49, 64, 69 enthalten keine Seitenüberschriften, ND 74 abweichende.
Die Zuordnung erfolgte sinngemäß.

Die Zusatzauswahlen ND 69/Z und ND 74 entstammen der Wochenendbeilage
"ND am Wochenende".

In ND 64 ist die Beilage "Die gebildete Nation" in den Hauptfile integriert (+ 8 Seiten).

1) incl. Theater-, Rundfunk-/Fernseh-Programme und Roman

2) ND 64: mit Berlin-Teil

2.5. Listen der aufgenommenen Seiten

Die folgenden Listen verzeichnen sämtliche nach dem statistischen Auswahlverfahren aufgenommenen Seiten des BZK 1, also jeweils die Modellmengen ohne Zusatzmengen. Es handelt sich um genau die Auswahl, die Grundlage der Microfiche-Veröffentlichung des BZK 1 ist, mit Ausnahme der Liste des Jahrgangs WE 49; dort sind die mit Z gekennzeichneten Seiten nicht in die Modellmenge aufgenommen worden.

Da den Seiten das genaue Datum und die Artikelnummer zugeordnet sind, kann jeder im BZK 1 gespeicherte Artikel und jeder daraus extrahierte Beleg datiert werden.

Die Artikelnummern in der rechten Spalte vermerken jeweils die niedrigste und höchste Artikelnummer auf der betreffenden Seite. Dazwischen liegende Nummern müssen nicht unbedingt besetzt sein. Beim Übergang von einem Blatt zum nächsten bleiben aus organisatorischen Gründen ohnehin meist Nummern frei.

Die Listen sind durchgehend chronologisch geordnet. Die Artikelnumerierung und (selten) die Blattnumerierung können springen. Auf dem Magnetband und in der Microfiche-Veröffentlichung (Gruppe Klartexte) liegen die Artikel dagegen in strikt numerisch aufsteigender Folge.

Die Numerierung in der linken Spalte bezieht sich auf die ausgewählten Seiten in der Textvorlage, die im IdS blattweise durchnummeriert, archiviert und gebunden wurde. Wer sich für den Originalartikel, seine Aufmachung und Platzierung auf der Seite etc. interessiert, kann über die Blattnummer am einfachsten darauf zurückgreifen und ggf. dazu eine Fotokopie/Rückvergrößerung beim IdS bestellen.

Z	1	06.01.	Do	5	2001-2023
	1	11.01.	Di	1	0030-0048
Z	2	22.01.	Sa	4	2055-2097
	2	27.01.	Do	2	0105-0118
Z	3	01.02.	Di	3	2125-2130
	3	12.02.	Sa	3	0135-0141
Z	4	17.02.	Do	2	2147-2162
	4	22.02.	Di	4	0170-0201
Z	5	05.03.	Sa	1	2170-2189
	5	10.03.	Do	5	0210-0234
Z	6	15.03.	Di	6	2200-2229
	6	26.03.	Sa	6	0240-0269
Z	7	31.03.	Do	4	2240-2280
	7	05.04.	Di	1	0275-0300
Z	8	16.04.	Sa	5	2290-2297
	8	20.04.	Mi	2	0306-0318
Z	9	26.04.	Di	2	2300-2313
	9	07.05.	So	3	0325-0330
Z	10	12.05.	Do	6	2320-2347
	10	17.05.	Di	4	0336-0364
Z	11	28.05.	Sa	7	2355-2376
	11	02.06.	Do	5	0370-0396
Z	12	08.06.	Mi	1	2385-2412
	12	18.06.	Sa	6	0401-0420
Z	13	23.06.	Do	6	2420-2448
	13	28.06.	Di	1	0426-0446
	14	06.07.	Mi	8	0452-0480
	15	14.07.	Do	5	0486-0512
	16	22.07.	Fr	2	0518-0527
	17	30.07.	Sa	3	0535-0542
	18	30.07.	Sa	5	0548-0569
	19	01.08.	Mo	4	0575-0603
	20	09.08.	Di	5	0610-0632
	21	17.08.	Mi	6	0637-0647
	22	25.08.	Do	1	0655-0678
	23	02.09.	Fr	2	0685-0697
	24	10.09.	Sa	9	0705-0710
	25	12.09.	Mo	3	0716-0724
	26	20.09.	Di	4	0730-0755
	27	28.09.	Mi	5	0761-0778
	28	06.10.	Do	6	0785-0807
	29	14.10.	Fr	1	0815-0837
	30	22.10.	Sa	4	0843-0867
	31	22.10.	Sa	8	0868-0887
	32	24.10.	Mo	1	0893-0912
	33	01.11.	Di	2	0917-0927
	34	09.11.	Mi	4	0933-0958
	35	17.11.	Do	4	0959-0965
	36	25.11.	Fr	3	0970-0979
	37	03.12.	Sa	7	0983-0997
	38	05.12.	Mo	6	1003-1017
	39	13.12.	Di	3	1023-1031
	40	21.12.	Mi	8	1037-1063
	41	29.12.	Do	2	1070-1081

WELT 1954/1 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
1	02.01.	Sa	1	1- 15
2	02.01.	Sa	11	16
3	02.01.	Sa	16	17- 28
4	11.01.	Mo	2	29- 44
5	19.01.	Di	3	45- 59
6	19.01.	Di	5	60- 71
7	27.01.	Mi	4	72- 89
8	04.02.	Do	5	101-116
9	12.02.	Fr	6	117-124
10	20.02.	Sa	7	125-130
11	20.02.	Sa	12	131-139
12	20.02.	Sa	17	140
13	22.02.	Mo	8	141-154
14	02.03.	Di	1	201-218
15	10.03.	Mi	2	219-239
16	10.03.	Mi	9	240-248
17	18.03.	Do	3	249-264
18	26.03.	Fr	4	265-271
19	26.03.	Fr	6	272-292
20	03.04.	Sa	5	301-308
21	03.04.	Sa	13	309-313
22	03.04.	Sa	22	314
23	05.04.	Mo	6	315-343
24	13.04.	Di	7	344-354
25	21.04.	Mi	8	355-371
26	29.04.	Do	1	372-392
27	07.05.	Fr	2	401-421
28	15.05.	Sa	3	422-435
29	15.05.	Sa	14	436-439
30	15.05.	Sa	30	440-448
31	17.05.	Mo	4	449-458
32	17.05.	Mo	7	459-476
33	25.05.	Di	5	477-482
34	02.06.	Mi	6	501-518
35	02.06.	Mi	10	519-534
36	10.06.	Do	7	535-538
37	17.06.	Do	8	539-542
38	17.06.	Do	9	543-551
39	26.06.	Sa	1	552-569
40	26.06.	Sa	15	570-574
41	26.06.	Sa	20	575
42	28.06.	Mo	2	576-593
43	06.07.	Di	3	601-614
44	06.07.	Di	8	615-627
45	14.07.	Mi	4	628-638
46	22.07.	Do	5	639-647
47	30.07.	Fr	6	648-666
48	07.08.	Sa	7	701
49	07.08.	Sa	16	702
50	07.08.	Sa	21	703-709
51	09.08.	Mo	8	710-722
52	17.08.	Di	1	723-742
53	25.08.	Mi	2	743-761
54	02.09.	Do	3	801-806

WELT 1954/2	IDS-BLATT	DATUM	TAG	SEITE	ART:-NR.
55		10.09.	Fr	1	807- 824
56		10.09.	Fr	4	825- 832
57		18.09.	Sa	5	833
58		18.09.	Sa	17	834
59		18.09.	Sa	22	835- 839
60		20.09.	Mo	6	840- 852
61		28.09.	Di	7	853- 862
62		06.10.	Mi	8	901- 920
63		14.10.	Do	1	921- 938
64		22.10.	Fr	2	939- 956
65		30.10.	Sa	3	957- 961
66		30.10.	Sa	12	962- 964
67		30.10.	Sa	23	965
68		01.11.	Mo	4	1001-1009
69		09.11.	Di	5	1010-1023
70		17.11.	Mi	6	1024-1027
71		25.11.	Do	7	1028-1030
72		25.11.	Do	10	1031-1049
73		03.12.	Fr	8	1101-1115
74		11.12.	Sa	1	1116-1134
75		11.12.	Sa	13	1135-1137
76		11.12.	Sa	24	1138
77		13.12.	Mo	2	1139-1159
78		13.12.	Mo	5	1160-1174
79		21.12.	Di	3	1175-1179
80		29.12.	Mi	4	1180-1197

WELT 1959/1	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1	02.01.	Fr	6	0001-0014
	2	10.01.	Sa	7	0020-0032
	3	10.01.	Sa	14	0040-0052
	4	12.01.	Mo	1	0060-0073
	5	12.01.	Mo	3	0080-0088
	6	20.01.	Di	8	0090-0104
	7	28.01.	Mi	11	0110-0125
	8	05.02.	Do	2	0130-0145
	9	13.02.	Fr	5	0150-0151
	10	21.02.	Sa	10	0160-0162
	11	21.02.	Sa	11	0170-0171
	12	21.02.	Sa	12	0180-0181
	13	21.02.	Sa	13	0190-0192
	14	23.02.	Mo	7	0200-0214
	15	03.03.	Di	1	0220-0232
	16	11.03.	Mi	2	0240-0258
	17	11.03.	Mi	3	0260-0267
	18	19.03.	Do	4	0270-0288
	19	19.03.	Do	11	0290-0293
	20	28.03.	Sa	5	0300-0310
	21	28.03.	Sa	15	0320-0321
	22	31.03.	Di	6	0330-0342
	23	08.04.	Mi	4	0350-0367
	24	08.04.	Mi	7	0370-0385
	25	08.04.	Mi	9	0390-0396
	26	16.04.	Do	8	0400-0407
	27	16.04.	Do	15	0410-0413
	28	24.04.	Fr	2	0420-0440
	29	24.04.	Fr	3	0450-0456
	30	24.04.	Fr	14	0460-0473
	31	02.05.	Sa	13	0480-0490
	32	04.05.	Mo	2	0500-0520
	33	04.05.	Mo	10	0530-0547
	34	12.05.	Di	5	0550-0557
	35	20.05.	Mi	1	0560-0575
	36	20.05.	Mi	6	0580-0590
	37	28.05.	Do	3	0600-0607
	38	28.05.	Do	9	0610-0613
	39	05.06.	Fr	4	0620-0637
	40	05.06.	Fr	8	0640-0649
	41	05.06.	Fr	11	0650-0651
	42	13.06.	Sa	2	0660-0672
	43	13.06.	Sa	16	0680-0693
	44	15.06.	Mo	3	0700-0707
	45	23.06.	Di	7	0710-0727
	46	23.06.	Di	10	0730-0747
	47	01.07.	Mi	5	0750-0759
	48	01.07.	Mi	12	0760-0769
	49	09.07.	Do	1	0770-0785
	50	09.07.	Do	6	0790-0796
	51	17.07.	Fr	9	0799-0809
	52	25.07.	Sa	4	0810-0824
	53	25.07.	Sa	8	0830-0844
	54	27.07.	Mo	6	0850-0854

WELT 1959/2 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
55	04.08.	Di	11	0860
56	04.08.	Di	12	0870-0888
57	12.08.	Mi	2	0890-0908
58	12.08.	Mi	7	0910-0932
59	20.08.	Do	5	0940-0946
60	20.08.	Do	10	0950-0958
61	20.08.	Do	12	1000-1010
62	28.08.	Fr	1	1020-1030
63	05.09.	Sa	3	1040-1045
64	05.09.	Sa	9	1050-1061
65	05.09.	Sa	13	1070-1080
66	07.09.	Mo	4	1090-1105
67	07.09.	Mo	8	1110-1126
68	15.09.	Di	1	1130-1147
69	23.09.	Mi	6	1150-1156
70	23.09.	Mi	8	1160-1164
71	01.10.	Do	2	1170-1184
72	09.10.	Fr	5	1190-1194
73	09.10.	Fr	10	1200-1201
74	09.10.	Fr	12	1210-1227
75	17.10.	Sa	1	1230-1246
76	17.10.	Sa	16	1250-1268
77	19.10.	Mo	6	1270-1284
78	27.10.	Di	8	1290-1302
79	27.10.	Di	10	1310-1311
80	04.11.	Mi	14	1320-1334
81	12.11.	Do	7	1340
82	20.11.	Fr	9	1350-1367
83	28.11.	Sa	10	1370-1373
84	28.11.	Sa	12	1380-1382
85	30.11.	Mo	2	1390-1406
86	30.11.	Mo	5	1410-1419
87	30.11.	Mo	9	1420-1427
88	08.12.	Di	7	1430-1445
89	08.12.	Di	9	1450-1453
90	08.12.	Di	11	1460
91	16.12.	Mi	3	1470-1475
92	16.12.	Mi	4	1480-1497
93	24.12.	Do	3	1500-1507
94	24.12.	Do	4	1510-1523
95	28.12.	Mo	1	1530-1547

WELT 1964/1 IDS-BLATT	DATUM	TAG	SEITE	ART, -NR.
1	02.01.	Do	1	1- 18
2	02.01.	Do	2	19- 32
3	02.01.	Do	4	35- 51
Z 1	04.01.	Sa	I	1801
Z 2	04.01.	Sa	II	1802-1805
4	10.01.	Fr	3	52- 58
5	10.01.	Fr	3	59- 63
Z 3	16.01.	Do	I	1806-1817
6	18.01.	Sa	5	65- 66
7	18.01.	Sa	6	67- 69
8	20.01.	Mo	7	70- 86
9	20.01.	Mo	8	87- 96
10	20.01.	Mo	9	100- 116
11	28.01.	Di	10	155- 171
12	28.01.	Di	11	172- 174
13	28.01.	Di	12	175- 186
14	05.02.	Mi	13	234- 240
15	05.02.	Mi	14	241- 256
16	13.02.	Do	15	288- 293
17	13.02.	Do	16	294- 300
Z 4	15.02.	Sa	II	1821-1826
Z 5	15.02.	Sa	IV	1820
18	21.02.	Fr	1	339- 357
19	21.02.	Fr	2	358- 373
20	21.02.	Fr	3	374- 380
21	29.02.	Sa	3	413- 417
22	29.02.	Sa	4	418- 427
23	29.02.	Sa	11	428- 441
24	02.03.	Mo	1	1600-1615
25	02.03.	Mo	5	442- 446
26	02.03.	Mo	6	447- 453
27	10.03.	Di	7	454- 470
28	10.03.	Di	8	471- 482
29	10.03.	Di	12	483- 493
30	18.03.	Mi	9	495- 513
31	18.03.	Mi	10	514- 522
32	26.03.	Do	11	523- 528
33	26.03.	Do	12	529- 540
34	26.03.	Do	13	541- 552
35	03.04.	Fr	15	553- 562
36	03.04.	Fr	16	564
Z 6	04.04.	Sa	III	1828
Z 7	04.04.	Sa	VII	1827
37	11.04.	Sa	17	582- 587
38	11.04.	Sa	18	588
39	11.04.	Sa	31	1616-1618
40	11.04.	Sa	32	1619-1620
41	13.04.	Mo	1	589- 605
42	13.04.	Mo	13	606- 621
43	13.04.	Mo	14	622- 632
44	21.04.	Di	19	633- 638
				648- 649
45	21.04.	Di	20	635, 700
46	29.04.	Mi	21	637, 639

WELT 1964/2 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
47	29.04.	Mi	22	640- 642
Z 8	07.05.	Do	11	1829-1832
48	15.05.	Fr	3	681- 687
49	15.05.	Fr	4	688- 698
50	15.05.	Fr	13	699- 707
Z 9	16.05.	Sa	IV	1834-1846
Z10	16.05.	Sa	X	1833
51	23.05.	Sa	5	721- 728
52	23.05.	Sa	6	708- 720
53	23.05.	Sa	48	1621
54	23.05.	Sa	57	1622-1623
55	25.05.	Mo	7	729- 750
56	25.05.	Mo	8	751- 762
57	02.06.	Di	5	764- 776
58	02.06.	Di	9	777- 792
59	02.06.	Di	10	793- 808
Z11	04.06.	Do	V	1847
60	10.06.	Mi	11	809- 829
61	10.06.	Mi	12	830- 843
62	17.06.	Mi	2	1624-1640
63	17.06.	Mi	13	844- 853
64	17.06.	Mi	14	854- 862
65	26.06.	Fr	1	1680-1696
66	26.06.	Fr	15	863- 864
67	26.06.	Fr	16	865- 870
68	04.07.	Sa	17	872
69	04.07.	Sa	18	873- 877
70	04.07.	Sa	26	1641-1644
71	04.07.	Sa	27	1645-1647
Z12	04.07.	Sa	V	1848-1852
72	06.07.	Mo	3	878- 884
73	06.07.	Mo	4	885- 904
74	14.07.	Di	5	905- 913
75	14.07.	Di	6	915- 924
76	22.07.	Mi	6	925- 930
77	22.07.	Mi	7	931- 949
78	22.07.	Mi	8	950- 959
79	30.07.	Do	9	960- 964
80	30.07.	Do	10	965
81	07.08.	Fr	5	1648-1652
82	07.08.	Fr	13	966- 985
83	07.08.	Fr	14	986- 991
84	15.08.	Sa	21	992
85	15.08.	Sa	22	993
Z13	15.08.	Sa	VI	1854-1855
Z14	15.08.	Sa	XVI	1853
86	17.08.	Mo	11	994- 995
87	17.08.	Mo	12	996-1012
88	25.08.	Di	1	1013-1031
89	25.08.	Di	2	1032-1049
90	25.08.	Di	8	1050-1064
Z15	27.08.	Do	III	1856-1862
91	02.09.	Mi	1	1065-1080
92	02.09.	Mi	2	1081-1097

WELT 1964/3 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
93	10.09.	Do	3	1098-1102
94	10.09.	Do	4	1103-1115
95	10.09.	Do	7	1116-1131
96	18.09.	Fr	5	1132-1142
97	18.09.	Fr	6	1143-1150
98	26.09.	Sa	23	1151-1152
99	26.09.	Sa	25	1153-1159
100	28.09.	Mo	7	1160-1177
101	28.09.	Mo	8	1178-1192
Z16	03.10.	Sa	I	1864-1867
Z17	03.10.	Sa	XIX	1863
102	06.10.	Di	9	1193-1208
103	06.10.	Di	10	1210-1229
104	14.10.	Mi	2	1700-1716
105	14.10.	Mi	11	1230-1242
106	14.10.	Mi	17	1243-1251
107	22.10.	Do	13	1256-1257
108	22.10.	Do	14	1258-1264
109	22.10.	Do	17	1653-1661
110	30.10.	Fr	14	1265-1274
111	30.10.	Fr	15	1275-1287
112	30.10.	Fr	16	1288-1302
113	07.11.	Sa	9	1303-1323
114	07.11.	Sa	19	1324-1328
115	07.11.	Sa	20	1329-1337
116	09.11.	Mo	1	1338-1354
117	09.11.	Mo	2	1355-1368
118	09.11.	Mo	10	1369-1376
Z18	14.11.	Sa	IV	1869-1879
Z19	14.11.	Sa	XXII	1868
119	17.11.	Di	3	1377-1384
120	17.11.	Di	4	1385-1400
121	25.11.	Mi	5	1401-1412
122	25.11.	Mi	6	1413-1424
123	25.11.	Mi	9	1662-1677
124	03.12.	Do	16	1425-1435
125	03.12.	Do	17	1436-1441
				1491-1498
Z20	03.12.	Do	IV	1880-1891
126	11.12.	Fr	7	1442-1460
127	11.12.	Fr	8	1461-1472
128	19.12.	Sa	14	1473-1484
130	19.12.	Sa	16	1504-1522
129	19.12.	Sa	19	1485-1490
131	19.12.	Sa	29	1678
132	21.12.	Mo	9	1523-1539
133	21.12.	Mo	10	1540-1554
134	29.12.	Di	11	1555-1570
135	29.12.	Di	12	1571-1588

WELT 1969/I	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
1		04.01.	Sa	1	201- 214
2		04.01.	Sa	2	217- 225
3		04.01.	Sa	3	228- 234
4		06.01.	Mo	2	238- 245
5		06.01.	Mo	3	251- 259
6		06.01.	Mo	4	264- 283
7		14.01.	Di	3	288- 296
8		14.01.	Di	4	301- 308
9		14.01.	Di	5	312- 325
10		22.01.	Mi	4	334- 352
11		22.01.	Mi	5	360- 368
12		22.01.	Mi	6	372- 378
13		30.01.	Do	5	382- 395
14		30.01.	Do	6	401- 403
15		30.01.	Do	7	406- 413
16		07.02.	Fr	6	417- 418
17		07.02.	Fr	7	421- 426
18		07.02.	Fr	8	430- 435
19		15.02.	Sa	7	445- 449
20		15.02.	Sa	8	465- 475
21		15.02.	Sa	9	480- 493
22		15.02.	Sa	22	500- 501
23		17.02.	Mo	8	508- 520
24		17.02.	Mo	9	525- 530
25		17.02.	Mo	10	535- 551
26		25.02.	Di	9	557- 562
27		25.02.	Di	10	567
28		25.02.	Di	11	571- 588
29		05.03.	Mi	10	593- 597
30		05.03.	Mi	11	601- 602
31		05.03.	Mi	12	605
32		13.03.	Do	11	610- 625
33		13.03.	Do	12	630- 639
34		13.03.	Do	13	645- 653
35		21.03.	Fr	12	658- 669
36		21.03.	Fr	13	675- 678
37		21.03.	Fr	14	683- 697
38		21.03.	Fr	21	700- 708
39		29.03.	Sa	15	712
40		29.03.	Sa	33	717- 719
41		29.03.	Sa	36	724- 728
42		31.03.	Mo	14	732
43		31.03.	Mo	15	736- 740
44		31.03.	Mo	16	745- 753
45		08.04.	Di	15	758- 770
46		08.04.	Di	16	775- 792
47		08.04.	Di	17	797- 806
48		16.04.	Mi	16	811- 814
49		16.04.	Mi	17	819- 827
50		16.04.	Mi	18	832- 836
51		24.04.	Do	17	841- 844
52		24.04.	Do	18	849- 853
53		24.04.	Do	19	858- 860
54		24.04.	Do	22	865- 873

WELT 1969/2 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
55	02.05.	Fr	18	878- 882
56	02.05.	Fr	19	887- 890
57	02.05.	Fr	20	895
60	10.05.	Sa	1	940- 951
58	10.05.	Sa	19	900- 912
59	10.05.	Sa	20	917- 931
62	12.05.	Mo	1	979- 991
63	12.05.	Mo	2	996-1003
61	12.05.	Mo	20	956- 974
64	20.05.	Di	1	1008-1020
65	20.05.	Di	2	1025-1032
66	20.05.	Di	3	1037-1040
67	28.05.	Mi	2	1045-1052
68	28.05.	Mi	3	1057-1059
69	28.05.	Mi	4	1064-1079
70	05.06.	Do	3	1084-1088
71	05.06.	Do	4	1093-1113
72	05.06.	Do	5	1118-1125
73	05.06.	Do	21	1130-1136
74	13.06.	Fr	4	1141-1157
75	13.06.	Fr	5	1162-1169
76	13.06.	Fr	6	1174-1184
77	21.06.	Sa	5	1189-1195
78	21.06.	Sa	6	1199-1207
79	21.06.	Sa	7	1212-1216
80	23.06.	Mo	6	1221-1231
81	23.06.	Mo	7	1236-1245
82	23.06.	Mo	8	1250-1256
83	01.07.	Di	7	1261
84	01.07.	Di	8	1264-1265
85	01.07.	Di	9	1270-1276
86	01.07.	Di	22	1281-1299
87	09.07.	Mi	8	1304-1307
88	09.07.	Mi	9	1310
89	09.07.	Mi	10	1315-1321
90	17.07.	Do	9	1326-1341
91	17.07.	Do	10	1345-1356
92	17.07.	Do	11	1361-1366
93	25.07.	Fr	10	1371
94	25.07.	Fr	11	1380-1395
95	25.07.	Fr	12	1400-1407
96	02.08.	Sa	22	1412-1414
97	02.08.	Sa	23	1420-1421
98	02.08.	Sa	24	1430-1434
99	04.08.	Mo	12	1439-1456
100	04.08.	Mo	13	1460-1469
101	04.08.	Mo	14	1474-1485
102	12.08.	Di	13	1490-1492
103	12.08.	Di	14	1497-1508
104	12.08.	Di	15	1513-1517
105	20.08.	Mi	14	1522-1536
106	20.08.	Mi	15	1541-1543
107	20.08.	Mi	16	1548-1551
108	28.08.	Do	15	1556-1558

WELT 1969/3 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
109	28.08.	Do	16	1563-1572
110	28.08.	Do	17	1577-1584
111	05.09.	Fr	16	1590
112	05.09.	Fr	17	1595-1610
113	05.09.	Fr	18	1615-1616
114	05.09.	Fr	21	1620-1625
115	13.09.	Sa	17	1630-1634
116	13.09.	Sa	18	1639-1648
117	13.09.	Sa	19	1653-1655
118	15.09.	Mo	18	1660-1669
119	15.09.	Mo	19	1674-1679
120	15.09.	Mo	20	1684-1693
121	15.09.	Mo	21	1698-1703
124	23.09.	Di	1	1738-1749
122	23.06.	Di	19	1708-1714
123	23.06.	Di	20	1719-1733
126	01.10.	Mi	1	1772-1784
127	01.10.	Mi	2	1789-1796
125	01.10.	Mi	20	1754-1767
128	09.10.	Do	1	1801-1813
129	09.10.	Do	2	1818-1825
130	09.10.	Do	3	1830-1832
131	17.10.	Fr	21	1837-1848
132	17.10.	Fr	24	1853-1862
133	17.10.	Fr	25	1867-1873
134	17.10.	Fr	26	1878-1896
135	25.10.	Sa	25	1901-1908
136	25.10.	Sa	26	1913-1918
137	25.10.	Sa	27	1925
138	27.10.	Mo	4	1935-1950
139	27.10.	Mo	5	1955-1962
140	27.10.	Mo	6	1967-1978
143	04.11.	Di	7	2012-2021
141	04.11.	Di	17	1983-1992
142	04.11.	Di	18	1997-2007
144	12.11.	Mi	27	2026-2030
145	12.11.	Mi	28	2035-2051
146	12.11.	Mi	29	2056
147	20.11.	Do	22	2061-2066
148	20.11.	Do	23	2071-2081
149	20.11.	Do	24	2086
150	20.11.	Do	25	2091-2097
151	28.11.	Fr	8	2102-2106
152	28.11.	Fr	9	2111-2118
153	28.11.	Fr	23	2123-2129
154	06.12.	Sa	30	2134-2135
155	06.12.	Sa	31	2140-2144
156	06.12.	Sa	32	2149-2150
157	08.12.	Mo	10	2155-2163
158	08.12.	Mo	11	2170-2183
159	08.12.	Mo	12	2188-2196
160	16.12.	Di	11	2204-2218
161	16.12.	Di	12	2228-2231
162	16.12.	Di	13	2236-2245

WELT 1969/4	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	163	24.12.	Mi	13	2250-2254
	164	24.12.	Mi	14	2259-2265
	165	24.12.	Mi	15	2270-2275
	166	31.12.	Mi	13	2280-2283
	167	31.12.	Mi	14	2288-2294
	168	31.12.	Mi	16	2299-2308

WELT 1974/1	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
1		02.01.	Mi	1	0001-0011
2		02.01.	Mi	2	0016-0034
3		02.01.	Mi	3	0040-0042
4		10.01.	Do	2	0047-0062
5		10.01.	Do	3	0067-0069
6		10.01.	Do	4	0074-0084
7		18.01.	Fr	3	0085-0090
8		18.01.	Fr	4	0095-0105
9		18.01.	Fr	5	0110-0118
10		26.01.	Sa	32	0124
11		26.01.	Sa	33	0130-0131
12		26.01.	Sa	34	0136-0141
13		28.01.	Mo	5	0146-0153
14		28.01.	Mo	6	0158-0168
15		28.01.	Mo	7	0174-0186
16		05.02.	Di	6	0191-0201
17		05.02.	Di	7	0206-0215
18		05.02.	Di	8	0225-0240
19		13.02.	Mi	7	0245-0251
20		13.02.	Mi	8	0256-0261
21		13.02.	Mi	9	0266-0285
22		21.02.	Do	8	0290-0301
23		21.02.	Do	9	0305-0320
24		21.02.	Do	10	0325-0339
25		01.03.	Fr	9	0345-0351
26		01.03.	Fr	10	0356
27		01.03.	Fr	11	0360-0379
28		09.03.	Sa	10	0384-0393
29		09.03.	Sa	11	0398-0412
30		09.03.	Sa	12	0415-0424
31		11.03.	Mo	11	0429
32		11.03.	Mo	12	0434-0435
33		11.03.	Mo	13	0440-0453
34		19.03.	Di	12	0460-0468
35		19.03.	Di	13	0475-0478
36		19.03.	Di	14	0480-0485
37		27.03.	Mi	13	0490-0496
38		27.03.	Mi	14	0500-0512
39		27.03.	Mi	15	0520-0522
40		04.04.	Do	14	0530-0537
41		04.04.	Do	15	0540-0543
42		04.04.	Do	16	0550-0558
43		13.04.	Sa	24	0565-0568
44		13.04.	Sa	27	0575-0579
45		13.04.	Sa	28	0585-0590
46		20.04.	Sa	16	0600-0608
47		20.04.	Sa	17	0615-0617
48		20.04.	Sa	18	0620
49		22.04.	Mo	17	0625-0633
50		22.04.	Mo	18	0640-0653
51		22.04.	Mo	19	0660-0669
52		30.04.	Di	18	0675
53		30.04.	Di	19	0680-0683
54		30.04.	Di	20	0690-0697

WELT 1974/2 IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
55	08.05.	Mi	19	0700-0710
56	08.05.	Mi	1	0720-0729
57	16.05.	Do	1	0735-0739
58	16.05.	Do	2	0745-0758
59	16.05.	Do	3	0765-0774
60	24.05.	Fr	2	0780-0795
61	24.05.	Fr	3	0800-0813
62	24.05.	Fr	4	0820-0831
63	01.06.	Sa	29	0840-0842
64	01.06.	Sa	30	0850-0853
65	01.06.	Sa	31	0860-0864
66	04.06.	Di	4	0870-0881
67	04.06.	Di	5	0890-0902
68	04.06.	Di	6	0910-0925
69	11.06.	Di	5	0930-0943
70	11.06.	Di	6	0950-0959
71	11.06.	Di	7	0965
72	19.06.	Mi	6	0970-0978
73	19.06.	Mi	7	0985
74	19.06.	Mi	8	0990-0997
75	27.06.	Do	7	1000-1007
76	27.06.	Do	8	1010-1019
77	27.06.	Do	9	1025-1035
78	05.07.	Fr	8	1040-1051
79	05.07.	Fr	9	1060-1071
80	05.07.	Fr	10	1080-1089
81	13.07.	Sa	9	1095-1097
82	13.07.	Sa	10	1100-1107
83	13.07.	Sa	11	1115-1124
84	15.07.	Mo	10	1130-1140
85	15.07.	Mo	11	1145-1151
86	15.07.	Mo	12	1160-1166
87	23.07.	Di	11	1175
88	23.07.	Di	12	1180-1192
89	23.07.	Di	13	1200-1209
90	31.07.	Mi	12	1215-1225
91	31.07.	Mi	13(1)	1230
92	31.07.	Mi	14(2)	1235-1236
93	08.08.	Do	13	1245-1249
94	08.08.	Do	14	1255-1267
95	08.08.	Do	17	1275-1286
96	16.08.	Fr	14	1295-1309
97	16.08.	Fr	1	1315-1327
98	16.08.	Fr	2	1335-1351
99	24.08.	Sa	1	1360-1371
100	24.08.	Sa	2	1380-1395
101	24.08.	Sa	3	1400-1410
102	26.08.	Mo	2	1415-1431
103	26.08.	Mo	4	1440-1452
104	26.08.	Mo	5	1460-1468
105	03.09.	Di	6	1475-1483
106	03.09.	Di	7	1490-1494
107	03.09.	Di	8	1500-1506
108	11.09.	Mi	9	1515-1527

WELT 1974/3	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR
109		11.09.	Mi	10	1535-1546
110		19.09.	Do	11(III)	1555-1560
111		19.09.	Do	12(IV)	1565-1568
112		27.09.	Fr	13	1575-1584
113		27.09.	Fr	14	1590-1594
114		05.10.	Sa	25	1600
115		05.10.	Sa	26	1605-1606
116		07.10.	Mo	15	1615-1622
117		07.10.	Mo	16	1630-1646
118		15.10.	Di	16	1635-1659
119		15.10.	Di	17	1665-1669
120		15.10.	Di	18	1675-1682
121		23.10.	Mi	17	1690-1695
122		23.10.	Mi	18(I)	1700-1703
123		23.10.	Mi	19(II)	1710-1712
124		31.10.	Do	20(II)	1720-1728
125		31.10.	Do	21(III)	1735-1738
126		31.10.	Do	22(IV)	1745-1751
127		08.11.	Fr	15	1760-1765
128		08.11.	Fr	20	1770-1781
129		08.11.	Fr	21	1790-1797
130		16.11.	Sa	22	1805-1808
131		16.11.	Sa	23	1815
132		18.11.	Mo	1	1820-1828
133		18.11.	Mo	15	1835-1843
134		26.11.	Di	1	1850-1860
135		26.11.	Di	16	1865-1870
136		26.11.	Di	17	1875-1879
137		04.12.	Mi	22	1885-1886
138		04.12.	Mi	23	1890-1893
139		04.12.	Mi	24(I)	1900
140		12.12.	Do	20(IV)	1910-1915
141		12.12.	Do	21	1920-1928
142		20.12.	Fr	3	1935-1943
143		20.12.	Fr	4	1950-1961
144		20.12.	Fr	5	1970-1979
145		28.12.	Sa	15	1985-1991
146		28.12.	Sa	16	2000-2016
147		28.12.	Sa	17	2025-2026
148		30.12.	Mo	4	2035-2047
149		30.12.	Mo	5	2055-2062
150		30.12.	Mo	6	2070-2078

ND 1949	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1	01.01.	Sa	5	0001-0003
	2	11.01.	Di	2	0010-0027
	3	19.01.	Mi	3	0035-0045
	4	27.01.	Do	7	0050-0064
	5	04.02.	Fr	1	0070-0092
	6	12.02.	Sa	4	0100-0125
	7	13.02.	So	2	0130-0149
	8	22.02.	Di	3	0155-0162
	9	02.03.	Mi	4	0170-0174
	10	10.03.	Do	1	0180-0203
	11	18.03.	Fr	2	0210-0234
	12	26.03.	Sa	3	0240-0254
	13	27.03.	So	8	0260-0273
	14	05.04.	Di	4	0280-0298
	15	13.04.	Mi	5	0310-0323
	16	21.04.	Do	1	0330-0345
	17	29.04.	Fr	6	0350-0363
	18	07.05.	Sa	2	0370-0395
	19	08.05.	So	4	0400-0402
	20	17.05.	Di	5	0410-0426
	21	25.05.	Mi	1	0435-0456
	22	02.06.	Do	6	0465-0476
	23	10.06.	Fr	2	0485-0515
	24	18.06.	Sa	3	0520-0529
	25	19.06.	So	5	0535-0546
	26	28.06.	Di	6	0550-0568
	27	06.07.	Mi	1	0575-0602
	28	14.07.	Do	2	0610-0633
	29	22.07.	Fr	3	0640-0651
	30	30.07.	Sa	4	0660
	31	31.07.	So	6	0670-0689
	32	09.08.	Di	1	0700-0721
	33	17.08.	Mi	2	0730-0752
	34	25.08.	Do	3	0760-0769
	35	02.09.	Fr	4	0775-0781
	36	10.09.	Sa	5	0790-0811
	37	11.09.	So	1	0820-0835
	38	20.09.	Di	2	0840-0864
	39	28.09.	Mi	3	0870-0884
	40	06.10.	Do	5	0890-0893
	41	14.10.	Fr	1	0900-0928
	42	22.10.	Sa	6	0935-0956
	43	23.10.	So	2	0965-0982
	44	01.11.	Di	3	0990-1000
	45	09.11.	Mi	4	1010-1015
	46	18.11.	Fr	5	1020-1037
	47	25.11.	Fr	6	1045-1066
	48	03.12.	Sa	1	1075-1088
	49	04.12.	So	3	1100-1110
	50	13.12.	Di	4	1120-1122
	51	21.12.	Mi	5	1130
	52	29.12.	Do	6	1135-1150

ND 1954	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1/2	01.01.	Fr	1	1- 5
	3	10.01.	So	3	700- 702
	4	19.01.	Di	4	28- 33
	5	27.01.	Mi	5	34- 49
	6	04.02.	Do	7	50- 65
	7	12.02.	Fr	2	66- 77
	8	20.02.	Sa	3	78- 86
	9	21.02.	So	4	703- 715
	10	02.03.	Di	5	95- 106
	11	10.03.	Mi	6	107- 123
	12	18.03.	Do	1	124- 136
	13	26.03.	Fr	3	137- 143
	14	03.04.	Sa	4	144- 148
	15	04.04.	So	7	716- 719
	16	13.04.	Di	5	153- 170
	17	21.04.	Mi	6	171- 188
	18	29.04.	Do	6	630- 632
	19	07.05.	Fr	4	206- 211
	20	15.05.	Sa	5	212- 225
	21	16.05.	So	2	720- 733
	22	25.05.	Di	1	230- 241
	23	02.06.	Mi	2	242- 253
	24	10.06.	Do	3	254- 258
	25	18.06.	Fr	5	259- 274
	26	26.06.	Sa	8	275- 291
	27	27.06.	So	7	735- 740
	28	06.07.	Di	6	300- 317
	29	14.07.	Mi	1	318- 326
					321
	30	22.07.	Do	2	327- 341
	31	30.07.	Fr	6	342- 357
	32	07.08.	Sa	1	633- 643
	33	08.08.	So	2	741- 753
	34	17.08.	Di	3	372- 379
	35	25.08.	Mi	4	380- 386
	36	02.09.	Do	5	387- 406
	37	10.09.	Fr	1	407- 421
	38	18.09.	Sa	2	422- 435
	39	19.09.	So	3	755- 765
	40	28.09.	Di	4	448- 450
	41	06.10.	Mi	5	451- 470
	42	14.10.	Do	6	473- 492
	43	22.10.	Fr	1	493- 504
	44	30.10.	Sa	2	505- 513
	45	31.10.	So	4	766- 770
	46	09.11.	Di	8	522- 544
	47	17.11.	Mi	3	545- 552
	48	25.11.	Do	4	553- 562
	49	03.12.	Fr	5	563- 571
	50	11.12.	Sa	6	572- 590
	51	12.12.	So	1	771- 776
	52	21.12.	Di	2	610- 620
	53	29.12.	Mi	3	603- 604

ND 1959	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1	01.01.	Do	1	0001-0007
	2	09.01.	Fr	3	0010-0018
	3	17.01.	Sa	4	0020-0029
	4	18.01.	So	5	0035-0036
	5	26.01.	Mo	2	0040-0050
	6	03.02.	Di	2	0055-0065
	7	11.02.	Mi	6	0070-0079
	8	19.02.	Do	7	0085-0100
	9	27.02.	Fr	1	0105-0118
	10	07.03.	Sa	3	0120
	11	08.03.	So	4	0125-0130
	12	16.03.	Mo	1	0135-0144
	13	24.03.	Di	5	0150-0167
	14	01.04.	Mi	2	0170-0183
	15	09.04.	Do	8	0190-0202
	16	17.04.	Fr	6	0210-0214
	17	25.04.	Sa	1	0220-0227
	18	26.04.	So	3	0230-0233
	19	04.05.	Mo	2	0240-0251
	20	12.05.	Di	5	0255-0266
	21	20.05.	Mi	1	0270-0278
	22	28.05.	Do	7	0280-0282
	23	05.06.	Fr	4	0290-0304
	24	13.06.	Sa	6	0310-0319
	25	14.06.	So	1	0325-0334
	26	22.06.	Mo	3	0340-0346
	27	30.06.	Di	2	0350-0365
	28	08.07.	Mi	4	0370-0372
	29	08.07.	Mi	5	0380-0388
	30	16.07.	Do	6	0395-0398
	31	24.07.	Fr	5	0400-0405
	32	24.07.	Fr	8	0410-0427
	33	01.08.	Sa	4	0430-0438
	34	02.08.	So	6	0445-0459
	35	10.08.	Mo	1	0460-0471
	36	18.08.	Di	3	0480-0489
	37	26.08.	Mi	2	0495-0509
	38	03.09.	Do	4	0515-0517
	39	11.09.	Fr	4	0520-0529
	40	19.09.	Sa	2	0535-0538
	41	20.09.	So	3	0540
	42	27.09.	Mo	8	0555-0576
	43	06.10.	Di	1	0580-0584
	44	14.10.	Mi	3	0590
	45	22.10.	Do	2	0595-0606
	46	30.10.	Fr	7	0610-0626
	47	07.11.	Sa	5	0630-0633
	48	07.11.	Sa	6	0640-0646
	49	08.11.	So	3	0650
	50	16.11.	Mo	4	0655-0667
	51	24.11.	Di	4	0670-0675
	52	02.12.	Mi	1	0680-0685
	53	10.12.	Do	3	0690-0693
	54	18.12.	Fr	2	0700-0712
	55	25.12.	Fr	5	0720-0733
	56	28.12.	Mo	6	0740-0755

ND 1964	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1	02.01.	Do	1	1- 11
	2	10.01.	Fr	2	12- 25
	3	18.01.	Sa	3	26- 30
	4/5	20.01.	Mo	2	31- 49
	6	28.01.	Di	3	68- 72
	7	05.02.	Mi	4	74
	8	13.02.	Do	5	75- 84
	9	21.02.	Fr	6	85
	10	29.02.	Sa	7	89
	11	02.03.	Mo	3	90- 98
	12	10.03.	Di	4	99- 109
	13	18.03.	Mi	5	110- 116
	14	26.03.	Do	6	117- 127
	15	03.04.	Fr	7	128- 129
	16	11.04.	Sa	5	130- 153
	17	13.04.	Mo	4	154- 169
	18	21.04.	Di	5	170- 189
	19	29.04.	Mi	6	190- 194
	20	15.05.	Fr	1	195- 200
	21	23.05.	Sa	2	209- 222
	22	25.05.	Mo	3	223- 234
	23	02.06.	Di	6	235- 245
	24	10.06.	Mi	7	246- 272
	25	18.06.	Do	8	273- 288
	26	26.06.	Fr	1	289- 299
	27	04.07.	Sa	2	300- 317
	28	06.07.	Mo	4	348- 357
	29	14.07.	Di	1	358- 369
	30	22.07.	Mi	2	370- 377
	31	30.07.	Do	3	378- 391
	32	07.08.	Fr	8	392- 405
	33	15.08.	Sa	6	406- 420
	34	17.08.	Mo	1	428- 439
	35	25.08.	Di	2	440- 455
	36	02.09.	Mi	3	207
	37	10.09.	Do	4	472- 482
	38	18.09.	Fr	5	483
	39	26.09.	Sa	8	484- 499
	40	28.09.	Mo	6	500- 508
	41	06.10.	Di	3	509- 513
	42	06.10.	Di	10	325- 335
	43	22.10.	Do	2	527- 542
	44	30.10.	Fr	1	543- 557
	45	07.11.	Sa	7	558- 578
	46	09.11.	Mo	3	579- 594
	47	17.11.	Di	4	595- 600
	48	25.11.	Mi	5	601- 614
	49	03.12.	Do	8	615- 631
	50	11.12.	Fr	7	632- 647
	51	19.12.	Sa	6	648- 666
	52	21.12.	Mo	1	667- 683
	53	29.12.	Di	5	684- 686

ND 1969/1	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1	04.01.	Sa	1	1- 9
	2	06.01.	Mo	2	11- 26
	4	14.01.	Di	3	41- 42
	5	22.01.	Mi	4	45- 46
	6	30.01.	Do	5	48- 69
	7	07.02.	Fr	6	72- 78
	8	15.02.	Sa	7	80- 95
	9	17.02.	Mo	8	97
	10	25.02.	Di	1	101- 113
	E 1	25.02.	Di	4	900- 910
	12	05.03.	Mi	2	121- 132
	13	13.03.	Do	3	135- 142
	14	21.03.	Fr	4	146- 153
	15	29.03.	Sa	5	155- 174
	16	31.03.	Mo	6	177- 196
	17	08.04.	Di	7	201- 220
	E 2	08.04.	Di	8	915- 924
	E 3	16.04.	Mi	3	930- 933
	18	16.04.	Mi	8	222- 239
	19	24.04.	Do	1	243- 252
	20	02.05.	Fr	2	255- 264
	21	10.05.	Sa	3	265- 267
	22	15.05.	Mo	4	270- 280
	23	20.05.	Di	5	283- 288
	24	28.05.	Mi	6	290- 295
	25	05.06.	Do	7	298- 316
	E 4	13.06.	Fr	5	938- 944
	26	13.06.	Fr	8	320- 335
	27	21.06.	Sa	1	339- 351
	28	23.06.	Mo	2	354- 365
	29	01.07.	Di	3	368- 379
	E 5	09.07.	Mi	1	950- 959
	30	09.07.	Mi	4	383- 383
	32	17.07.	Do	5	412- 422
	33	25.07.	Fr	6	426- 432
	E 6	02.08.	Sa	6	965- 967
	34	02.08.	Sa	7	435- 448
	35	04.08.	Mo	8	451- 452
	36	12.08.	Di	1	455- 468
	37	20.08.	Mi	2	471- 481
	38	28.08.	Do	3	484- 489
	40	05.09.	Fr	4	497- 504
	E 7	05.09.	Fr	7	972- 993
	41	13.09.	Sa	5	507- 512
	42	15.09.	Mo	6	515
	43	23.09.	Di	7	518- 539
	44	01.10.	Mi	8	543- 549
	48	09.10.	Do	1	561- 569
	49	17.10.	Fr	2	572- 585
	50	25.10.	Sa	3	588
	51	27.10.	Mo	4	590- 601
	52	04.11.	Di	5	603- 615
	53	12.11.	Mi	6	618- 622
	E 8	20.11.	Do	2	995-1009

ND 1969/2	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	54	20.11.	Do	7	624- 649
	55	28.11.	Fr	8	652- 665
	56	06.12.	Sa	1	668- 677
	57	08.12.	Mo	2	680- 702
	59	16.12.	Di	3	726
	60	24.12.	Mi	4	728- 737
	61	31.12.	Do	5	741- 747

ND 1974/1	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	1	02.01.	Mi	1	0001-0012
	2	10.01.	Do	2	0018-0029
	3	18.01.	Fr	3	0035
	4	26.01.	Sa	4	0041-0047
	5	28.01.	Mo	5	0055-0079
	6	05.02.	Di	3	0085
	7	05.02.	Di	6	0091-0098
	8	13.02.	Mi	7	0104-0128
	9	21.02.	Do	4	0134-0145
	10	21.02.	Do	8	0151-0164
	11	01.03.	Fr	2	0170-0183
	12	09.03.	Sa	3	0189-0193
	13	11.03.	Mo	4	0198-0214
	14	19.03.	Di	5	0219-0233
	15	27.03.	Mi	6	0238-0242
	16	04.04.	Do	7	0247-0270
	17	13.04.	Sa	5	0295-0307
	18	20.04.	Sa	1	0312-0317
	19	22.04.	Mo	3	0322-0327
	20	30.04.	Di	4	0337-0350
	21	08.05.	Mi	3	0355
	22	08.05.	Mi	5	0360-0368
	23	16.05.	Do	6	0373-0376
	24	24.05.	Fr	4	0381-0392
	25	24.05.	Fr	7	0396-0416
	26	01.06.	Sa	8	0426-0436
	27	04.06.	Di	1	0441-0449
	28	11.06.	Di	2	0454-0471
	29	19.06.	Mi	4	0476-0489
	30	27.06.	Do	5	0494-0511
	31	05.07.	Fr	6	0516-0520
	32	13.07.	Sa	7	0525-0549
	33	15.07.	Mo	8	0556-0574
	34	23.07.	Di	1	0579-0584
	35	31.07.	Mi	2	0589-0609
	36	08.08.	Do	3	0614-0622
	37	16.08.	Fr	5	0627-0634
	38	24.08.	Sa	6	0639-0643
	39	26.08.	Mo	2	0648-0670
	40	26.08.	Mo	7	0676-0689
	41	03.09.	Di	6	0695-0699
	42	03.09.	Di	8	0705-0725
	43	11.09.	Mi	1	0731-0741
	44	19.09.	Do	2	0747-0764
	45	27.09.	Fr	5	0770-0781
	46	05.10.	Sa	4	0787-0794
	47	07.10.	Mo	6	0800-0812
	48	15.10.	Di	7	0818-0842
	49	23.10.	Mi	8	0852-0873
	50	31.10.	Do	1	0879-0888
	51	08.11.	Fr	2	08 -0909
	52	08.11.	Fr	8	0915-0928
	53	16.11.	Sa	3	0935-0946
	54	18.11.	Mo	1	0951-0963

ND 1974/2	IDS-BLATT	DATUM	TAG	SEITE	ART.-NR.
	55	18.11.	Mo	4	0968-0982
	56	26.11.	Di	5	0988
	57	04.12.	Mi	7	0995-1023
	58	12.12.	Do	8	1030-1049
	59	20.12.	Fr	1	1055-1064
	60	28.12.	Sa	2	1070-1088
	61	30.12.	Mo	3	1095-1106

2.6 Statistische Angaben

Text	1 Artik.	2 Sätze	3 Zeilen	4 token	5 types	6 type/ token Rel.	7 Satz- länge	8 Artikel- länge	9 Text- läng. Fakt.
D49	820	13498	31561	206186	32077	6,43	15,28	251,44	0,79
D54	591	12931	34255	207624	27869	7,45	16,06	351,31	1,24
D59	522	13044	31914	208003	29231	7,12	15,95	398,47	1,41
D64	630	12625	33285	193844	31288	6,20	15,35	307,69	1,96
D69	649	10775	28390	173554	26818	6,48	16,11	267,42	2,39
D74	795	11135	28390	179043	28022	6,39	16,08	225,21	2,03
D..	4007	74008	187795	1168254	175305	6,66	15,79	291,55	1,60
W49	768	12251	27073	163359	33383	4,9	13,33	212,71	1,26
W54	901	19211	46626	258281	42794	6,04	13,44	286,66	0,80
W59	1063	22272	48818	292774	47409	6,18	13,15	275,42	0,71
W64	1428	26905	69667	379429	58297	6,51	14,10	265,71	0,51
W69	1411	31246	77838	414466	64179	6,46	13,26	293,74	0,42
W74	1262	26869	62213	362957	59662	6,42	13,51	287,60	0,49
W..	6833	138754	332235	1871266	305724	6,12	13,49	273,86	0,62
Alle	10840	212762	520030	3039520	481029	6,32	14,29	280,40	

- Spalte 1 = Zahl der Artikel
 Spalte 2 = Zahl der Sätze
 Spalte 3 = Zahl der Zeilen im Ausdruck bzw. Microfiche
 Spalte 4 = Zahl der lfd. Wörter (tokens), u = u Textlänge
 Spalte 5 = Zahl der verschiedenen Wortformen (types) (ohne Sonderzeichen und Transkriptionszeichen, incl. Zahlen), = Vokabular
 Spalte 6 = type-token-Relation (tokens/types)
 Spalte 7 = Durchschnittliche Satzlänge (tokens/Sätze)
 Spalte 8 = Durchschnittliche Artikellänge (tokens/Artikel)
 Spalte 9 = Textlängenfaktor (tokens WELT/tokens ND)
 (tokens ND/tokens WELT)

3. ZUR CODIERUNG DES BZK I

3.1. Allgemeines zur Textcodierung

Unter Textcodierung verstehe ich zunächst jede Art von absichtlicher Veränderung an Texten, die man vornimmt, wenn man Texte auf Datenträger übertragen und maschinenverfügbar machen will.

Auch "Text" ist sehr allgemein zu verstehen: gemeint sind damit nicht nur die Buchstaben und Zeilen, die sich zu Wörtern, Sätzen und Text im linguistischen Sinne verbinden (das wäre sehr einfach), sondern auch alles andere, was sich z.B. auf einer Zeitungsseite befindet, z.B. Bilder (Fotos), Tabellen und Graphiken, Spartenüberschriften und Impressum, Werbezeichnungen, Firmenembleme, Warenzeichen und Lottozahlen. Es gehören dazu auch die besonderen Mittel, die eine Zeitung anwendet, um Texte zu gliedern, Textteile hervorzuheben oder in Beziehung zu anderen zu setzen, wie unterschiedliche Schriftgrade (z.B. bei Überschriften) oder Schriftstärken, mehrspaltiger Satz, Einrückungen, Sperrungen, Striche, Punkte und Pfeile, Symbolzeichen der verschiedensten Art (z.B. für Telefon, Paragraph, Prozent, Dollar usw.).

Gegenüber dieser Fülle an Möglichkeiten mutet der Zeichenvorrat in DV-Anlagen und die Möglichkeiten ihn einzusetzen geradezu ärmlich an: Auf den früheren Lochkartengeräten waren gerade 48 Zeichen, auf den späteren 64 Zeichen verfügbar; ein moderner Bildschirm mit Groß-/Kleinschreibung verfügt über 84 Zeichen, die normalen Schnelldrucker-Ketten über 126 Zeichen. Hinzu kommt, daß in der Textvorlage nebeneinander Geordnetes (wie Mehrspaltensatz, Tabellen, Graphiken oder Zeichnungen mit integriertem Text) bei der Erfassung in eine lineare Ordnung gebracht werden muß.

Eine Veränderung oder Bearbeitung ist also in jedem Falle unerlässlich. Die Frage ist, in welchem Sinne diese Veränderung vorgenommen werden soll.

Eine Textdokumentation, die für möglichst viele verschiedene, auch vor-

her nicht bekannte Zwecke dienen soll, wird vor allem der Forderung nach möglichst weitgehender dokumentarischer Treue folgen, also dem Grundsatz: so wenig abändern wie möglich. Das impliziert die Forderung nach Vollständigkeit und nach zeichengetreuer Umsetzung und nach Verzicht auf Zutaten.

Dagegen wird der Benutzer, hier also der auswertende Linguist, auch wenn er die "ganze" Zeitung als Textgrundlage wünscht, "Vollständigkeit" nicht so umfassend verstehen wie der Dokumentar: Gleichbleibende Teile wie der Zeitungskopf, Spartenüberschriften, Impressum, Preisangaben der Zeitung etc. interessieren ihn nicht, selten auch Zahlenkolonnen, wie sie in den täglichen Börsenübersichten vorkommen. Dafür interessieren ihn um so mehr linguistische Merkmale, die der Originaltext in aller Regel keineswegs in der gewünschten Eindeutigkeit oder Vollständigkeit liefert, z.B. eine Markierung der Satzgrenzen und der Satzgliederung oder z.B. eine Beschränkung der Großschreibung auf Substantive und Anredenpronomen; auch wünscht er vielleicht Überschriften von Normaltext oder deutschen Text von fremdsprachigem zu unterscheiden oder die Anführungszeichen funktional zu differenzieren. Das alles mußte bei der Erfassung durch Auslassungen oder Einfügungen unterschiedlicher Art geregelt werden. Bei der Formulierung dieser Codierregeln tritt erfahrungsgemäß eine Schwierigkeit auf: Vieles was dem später auswertenden und deshalb vorher regelformulierenden Linguisten ganz klar zu sein scheint, ist so klar in der Praxis nicht, und vor allem nicht den Schreibkräften, die die Regeln anzuwenden haben (z.B.: was ist ein Name? was ist ein Satzschluß?). Je mehr Ausnahmeregeln oder Interpretationshilfen eine Codierregel nach sich zieht, um so mehr spricht dafür, auf sie zu verzichten und lieber genau so zu schreiben, wie es in der Textvorlage steht. Auch bei der Codierung der Bonner Zeitungstexte ist die anfängliche Codierungseuphorie einer zunehmenden Skepsis gewichen.

Grundsätzlich muß aber für jede Codierung gelten: Sie muß als solche formal erkennbar sein und sie muß reversibel sein. Eine Codierung, die den ursprünglichen Textzustand nicht wieder herzustellen erlaubt, verfälscht ihn und macht darauf bezogene Auswertungen unmöglich. Man muß gute Gründe haben, um dies in Kauf zu nehmen.

Sind die Forderungen des Dokumentars nach Vorlagentreue und des auswertenden Linguisten nach erleichterter Auswertbarkeit durch Codierung schon oft schwer vereinbar, so tritt als dritte Forderung die des Datenverarbeiters nach formaler Konsequenz hinzu: Gleiches habe auch formal gleich zu sein (= identische Zeichenketten), Zusammengehöriges muß auch zusammen geschrieben sein (kein Leeranschlag im "Wortinneren"); zwischen Wort und Nichtwort, Satz und Nichtsatz, Text und Nichttext ist eindeutig zu unterscheiden. Diese Forderung führt unter Umständen zu Veränderungen am und im Wort, an Satzzeichen, am Text überhaupt, die einer der beiden obengenannten Forderungen oder beiden widersprechen. Obwohl die Benutzung z.B. alphabetischer Register ohne Zweifel die Berücksichtigung bestimmter Sortierkriterien - und damit bestimmter Codierungen - nahelegt, ist auch hier Skepsis am Platz: Es hat keinen Zweck, aus Gründen der Datenverarbeitung den Texten eine Einheitlichkeit überzustülpen, die diese nun einmal nicht haben. Schreibvarianten und Sortierprobleme wird es immer geben. Perfekte Register sind nur um den Preis starker Eingriffe in den Text und/oder sehr hohen Codieraufwandes zu haben.

Eine letzte Forderung, die in der Praxis sogar von ausschlaggebendem Gewicht sein kann, kollidiert möglicherweise mit allen drei vorgenannten: die Forderung nämlich, daß auf ein bestimmtes, schon vorhandenes und lange praktiziertes Codiersystem Rücksicht zu nehmen ist und daß neue Texte, auch um eine einheitliche maschinelle Auswertung zu gewährleisten, sich diesem System anzupassen haben, - Forderungen eins bis drei hin oder her.

So galt z.B. für das Mannheimer Korpus von Anfang an die Regel: Jeder Satz wird durch Punkt zwischen blanks (= Leeranschlag) geschlossen. Nach Fragezeichen, Ausrufezeichen wird zusätzlich ein Satzschlußpunkt gesetzt. Wenn gar kein Satzschlußzeichen vorhanden ist, wird trotzdem Punkt zwischen blanks gesetzt. Grund für die strikte Satzschlußregelung war, daß der linguistische Ansatz für das Mannheimer Korpus 1 vor allem ein grammatisch-syntaktischer war: es sollten Sätze gefunden und analysiert werden. Da das Korpus ganz überwiegend aus literarischen Texten bestand, traten Probleme kaum auf: Die meisten Sätze schlossen ohnehin mit Punkt; die Codierung für Fragezeichen und Ausrufezeichen war re-

versibel, Sätze ohne Satzschlußpunkt kamen fast nur in Überschriften vor, und diese wiederum waren codiert, genauer: transkribiert.

Erst bei der Anwendung auf Zeitungstexte wurde die Problematik erkennbar: Bis zu 15 % der "Sätze" sind in der Textvorlage nicht durch Satzschlußpunkte abgeschlossen. Vor allem in Tabellen, Listen, Schaubildern, in zahlreichen Werbeanzeigen werden andere als die laut DUDEN vorgesehenen Mittel zur Gliederung von Äußerungen benutzt (z.B. solche des layout). Aber auch in Familien- und Kleinanzeigen, bei Aufrufen, Parolen, Slogans, in Bildbeischriften, Bilanzen und Bekanntmachungen wird vom Mittel des klassischen Satzschlußpunktes manchmal nur sparsam oder ungewöhnlicher Gebrauch gemacht. Hier wird die sonst unproblematische Codierregel für Satzschlüsse nicht nur für die Schreibkräfte problematisch, sondern auch, mangels "richtiger Sätze", für die spätere linguistische Auswertung. Vor allem aber verstößt sie gegen das Prinzip der Reversibilität: Man kann den DV-gespeicherten Texten nicht mehr ansehen, was in der Textvorlage gestanden hat.

Gleichwohl wurde bei der Anpassung des BZK an die Mannheimer Konventionen diese Regelung aus den erwähnten Gründen übernommen und angewandt. Uns schien dies vertretbar, da die ursprüngliche Bonner Version ja weiterhin zugänglich ist. Wer sich also für die Frage interessiert, wie häufig z.B. in den Bonner Zeitungstexten der Satzschlußpunkt wirklich vorkommt, kann die Bonner Version benutzen.

Im übrigen ist die auf das BZK angewandte Codierung, wie nicht anders zu erwarten, ein Kompromiß.

Als oberster Grundsatz galt allerdings: So textnah wie möglich. Gegenüber dem Gebot der Vollständigkeit wurden gewisse Abstriche gemacht: Zeitungskopf, Impressum, Spartenüberschriften und anderes stets Gleichbleibendes wird als Nicht-Text definiert und übergangen. Anderes, das im laufenden Text vorkommt, aber ohne erkennbares linguistisches Interesse ist, wie z.B. Zahlenkolonnen, wird ausgelassen, aber das Ausgelassene durch Hinweis gekennzeichnet. Auch Dinge, die durch Schrift nicht wiederzugeben sind, wie Graphiken, Firmenembleme, Zeichnungen, Fotos, Sym-

bolzeichen, werden ausgelassen, aber als Ausgelassenes gekennzeichnet. Drucktechnische Besonderheiten wie Halbfettdruck, Sperrdruck, Kursive o.ä. werden im BZK nicht markiert.

3.2 Arten von Codierungen

Wir unterscheiden

- a) Codierungen am Wort, die bei der Textzerlegung sortierrelevant werden:

z.B. Großschreibung am Satzanfang wird ausgeglichen; Großgeschriebene Adjektive als Teil von Eigennamen und Personen-Familiennamen erhalten Zusatzmarkierungen.

Hybridwörter aus Zahlen und Buchstaben werden mit Bindestrich durchgekoppelt, (7-1/2-t-Kipper statt 7 1/2 t Kipper) etc.

- b) Codierungen im Satz, die Sonderzeichen ihrer Funktion nach (syntaktische oder andere) vereindeutigen:

z.B. alle Zeichen mit syntaktischen Funktionen stehen zwischen blanks etc.

- c) Codierungen bestimmter Textstellen, die es erlauben, diese Textstellen gesondert anzusprechen oder zu übergehen. (Diese Codierungen heißen Transkriptionen):

z.B. Überschriften, fremdsprachige Textstellen usw. werden durch Einfügen von Transkriptionszeichen am Anfang und am Ende markiert.

- d) Codierungen an der Textgestalt

z.B. Überschriften werden durch Leerzeile vom Normaltext getrennt; Absätze werden durch Einrückung von 3 Leeranschlügen oder durch Leerzeile markiert.

Diese Codierungen sollen die Lesbarkeit des dv-gespeicherten Textes verbessern.

Die Forderung nach Reversibilität ist - mit Ausnahme der erwähnten Satzschluß-Regelung und drei weiteren sehr marginalen Punkten (siehe unter 3.3, Absatz 4)) berücksichtigt.

Abweichungen von der Textvorlage aus Gründen der Sortierung wurden nur sehr sparsam zugelassen und gegenüber der früheren Mannheimer Konvention noch weiter vermindert.

3.3. Aus den Codiervorschriften für das BZK 1

Vorbemerkung: Die folgenden Codierungen und Regeln gelten für die überarbeitete, den Mannheimer Konventionen weitgehend angepaßte Version des BZK 1. Für die Bonner Ausgangsversion gelten in den mit (*) (Sternchen) bezeichneten Fällen andere Regeln, die bei Benutzung der Bonner Version zu erfragen sind.

1. Drucktechnisches

Impressum und Seitenüberschriften werden generell übergangen, ebenso der Zeitungskopf auf der Titelseite.

Sperrung, Kursive, Halbfett o.ä. werden nicht gekennzeichnet (vgl. aber Überschriften), ebenso wenig diakritische Zeichen (Akzente u.ä.).

Versalienschreibung wird umgesetzt in Normalschreibung, Abkürzungen bleiben unverändert.

Zur Textgliederung werden, möglichst in Anlehnung an die Vorlage, Leerzeile und/oder Einrückung um 3 blanks verwendet.

Fette Punkte zur Textgliederung am Zeilenanfang werden durch Spiegelstriche zwischen blanks wiedergegeben.

Überschriften und Zwischenüberschriften werden vom Normaltext durch Leerzeile abgesetzt.

ORIGINAL
DIE WELT
GmbH

ABSCHRIFT
die Welt
GmbH

2. Transkriptionen(*)

a+	+a =	übersetzte Textstelle(*)
c+	+c =	Bildbeischrift(*)
f+	+f =	Fremdsprachliche Textstelle
m+	+m =	Mundartliche Textstelle
o+	+o =	Ostzitat in Westtext
q+	+q =	Quellenangabe: Ort, Tag, Agentur (am Kopf von Artikeln)
u+	+u =	Überschriften, Zwischenüberschriften
v+	+v =	Verfasserangaben (am Anfang oder Schluß von Artikeln)
w+	+w =	Westzitat in Osttext
x+	+x =	Nichtabdruckbare Symbolzeichen (z.B. x+ Durchmesser +x, x+ Pfund Sterling +x), nicht Schreibbares (x+ Foto +x, x+ Tabelle +x etc.) und Schreibkommentare (z.B. x+ Textstelle unlesbar +x)

y+ +y = Wegzulassendes
(z.B. Anzeigenschiffren, Zahlenkolonnen und Börsentabellen in Bilanzen etc.)

3. Behandlung von Zusammensetzungen und Abkürzungen

ORIGINAL	ABSCHRIFT
Mit-sich-geschehen-lassen	Mit-sich-geschehen-lassen
Wald- und Wiesen-Tee	Wald- und Wiesen-Tee
SPD - Fraktion	SPD-Fraktion
Dr.	Dr.(*)
Str.	Str.(*)
"Entwicklungs"-Völker	"Entwicklungs"-Völker(*)
(Wasser)Ball	(Wasser)Ball(*)
(Wasser-)Ball	(Wasser-)Ball(*)
CDU / CSU	CDU/CSU
1980 / 81	1980/81(aber: 1933 - 1945)
km/h	km/h
6 1/2 Prozenter	6-1/2-Prozenter

4. Zahlen und Zahlenverhältnisse; Mathematische Ausdrücke und Formeln

Dezimalzahlen gelten als ein Wort, werden also ohne Blank vor und nach dem Komma geschrieben.

Der angelsächsische Dezimalpunkt wird als Dezimalkomma geschrieben.

ORIGINAL	ABSCHRIFT
123,45	123,45
4.2	4,2
150 000	150000
1.250.000,25	1.250.000,25
11	1-1/2(*)
$\frac{7}{9}$	
9.30	9.30 Uhr

Mathematische Operatoren werden wie in der Textvorlage geschrieben.
Malzeichen: x*)

Formeln werden wie folgt geschrieben

10^5	10E5 (E steht für Exponent)(*)
N_2H_4	N2H4
$2N(NH_3)_2Cl_2$	((2N(NH3)2CL2))(*)

Eckige Klammern sind als nicht abdruckbare Zeichen durch runde Klammern zu ersetzen.(*)

Römische Zahlen werden in arabische Ziffern umgesetzt.

Als einziges Symbolzeichen ist das Zeichen % (Prozent) zulässig.(*). Alle anderen werden umschrieben und mit x+ ... +x transkribiert (siehe 2.).

5. Behandlung von Satzzeichen

Zwischen Satzzeichen und Wort (Zahl) bzw. Satzzeichen und Satzzeichen steht immer ein blank.

Das letzte Zeichen eines Satzes ist immer Punkt - auch dann, wenn in der Vorlage ein anderes oder gar kein Satzschlußzeichen steht. (*) Deshalb sind am Ende des Satzes ggf. folgende Umstellungen bzw. Punkt-Ergänzungen erforderlich.

ORIGINAL	ABSCHRIFT
.)) . } Punkt wird umgesetzt
! "	! " .
? "	? " .
)) . } Punkt wird ergänzt
***	*** .
mit folgender Großschreibung	mit folgender Großschreibung

Auch Buchtitel, Untertitel und Überschriften gelten als Satz und sind mit Punkt zu schließen. Danach folgt eine Leerzeile.

Ausnahme: Überschriften, die Teil des ersten Satzes eines Textes sind.

Beispiel:

ORIGINAL	ABSCHRIFT
WIR SIND DIE BESTEN	u+ wir sind die Besten +u
sagte Uwe gestern zu	sagte Uwe gestern zu
unserem Reporter	unserem Reporter .

Regelung bei Aufzählungen

Aufzählungsmerkmale gelten als eigenständige Sätze und bekommen einen Satzpunkt, wenn vollständige Sätze folgen, deren Anfang in der Vorlage groß geschrieben wird.

Abschnittsgliederungsmerkmale, die ohne Einleitung größere Passagen (z. B. mehrere Seiten) einleiten und daher eigentlich nicht als Aufzählung betrachtet werden können, werden (wenn sie nicht in den Satzzusammenhang eingehen) wie Aufzählungsmerkmale behandelt.

ORIGINAL	ABSCHRIFT
1)	1) .
2)	2) .
3)	3) .
a)	a) .
b)	b) .

6. Groß/Kleinschreibung

Am Satzanfang wird klein geschrieben (es sei denn, es stünde ein Substantiv, Name oder Anredepronomen dort).

Groß geschriebene Adjektive in Namen werden durch Apostroph plus Großschreibung gekennzeichnet. Ebenso wird bei Adjektiven verfahren, die von Städtenamen abgeleitet sind. (Versalien siehe obenb unter 1.)

ORIGINAL

Grüne Woche
Berliner Bevölkerung

ABSCHRIFT

'Grüne Woche
'Berliner Bevölkerung

7. Personen/Namen

Personen-Familiennamen werden mit vorgesetztem Stern gekennzeichnet:
Helmut *Schmidt(*)

8. Worttrennung

Worttrennung am Zeilenende ist nicht zulässig(*)

4. ZUR IDENTIFIKATION UND KLASSIFIKATION DER ZEITUNGSARTIKEL

4.1. Informationsklassen der IK-Datei

Die Dokumentationseinheit, d.i. die Datenmenge, die stellvertretend für ein Dokument in den Dokumentationsprozeß eingeht, ist in unserem Fall der einzelne Zeitungsartikel. Als Bestandteile von sogenannten Sammeldokumenten stellen die einzelnen Zeitungsartikel bibliographisch unselbständige Dokumentationseinheiten dar.

Der Begriff "Zeitungsartikel" hat in unserer Dokumentationspraxis einen relativ weiten Inhalt: Wir verstehen darunter auch z.B. eine Sammlung von Kurznachrichten unter einer gemeinsamen Überschrift (wie etwa "Kurz berichtet"), eine Sammlung von Kfz.-Anzeigen, eine Börsentabelle, eine Lotto-Gewinnliste, den "Pressespiegel", "Roman unterm Strich" usw.

Jeder Zeitungsartikel trägt eine Nummer und zusätzlich die Jahrgangssigle zur Kennzeichnung, aus welchem Zeitungsjahrgang er stammt.

Diese Artikelnummer und die Jahrgangssigle sind zugleich der Schlüssel oder die Adresse zu einem Satz von Informationen, die jedem Zeitungsartikel zugeordnet sind. Diese Informationen lassen sich in 3 Gruppen gliedern:

- Informationen zur Wiederauffindung und Datierung des Artikels

- Informationen zu seiner Herkunft (Autor/Agentur)
- Informationen zu Inhalt, Form und Aufmachung.

Alle diese Informationen sind, zum Teil verschlüsselt, in einem IK-Satz (IK = Information und Klassifikation) von jeweils 80 Zeichen Länge enthalten; alle IK-Sätze zusammen bilden die IK-Datei.

Sie ist folgendermaßen aufgebaut (vgl. das folgende Muster der IK-Datei):

Gruppe A

1. Korpussigle (BK = Bonner Korpus)
2. Satzkennung (IK = IK-Datei)
3. Artikelnummer 4-stellig
4. Angabe, ob der Artikel unvollständig ist (F = Fragment)
5. Angabe der Publikationsgattung (ZT = Tageszeitung)
6. Angabe, ob Ost- oder Westtext (O/W)
7. Aufnahmeprinzip (1 = statistisches; Modellmenge)
8. Zeitungsjahrgang (NDB = Neues Deutschland, Republik Ausgabe)
9. Datierung (69010602" Jg. 69, Monat: Januar, Tag: 6., Seite: 02)

Gruppe B

10. Angabe des Verfassers (falls voller Name: Sternchen vorangestellt)
11. Angabe der Agentur(en)

Gruppe C

12. Angabe des Sachgebiets (bis zu 3-fach-Kennzeichnung möglich)
13. Angabe der Intention (publizistisches Ziel), (2-fach-Kennzeichnung möglich)
14. Angabe der Artikelform (zeitungsspezifische Textsorte)
15. Angabe der Sparte (1 = Politik)
16. Angabe der Aufmachung (Länge und Breite des Artikels)

Auflösung der Jahrgangssiglen in Feld 8:

BG = Bonner General-Anzeiger (Stadttausgabe)

FR = Frankfurter Rundschau (Deutschlandausgabe)

ND = Neues Deutschland (A = Berliner Ausgabe, B = Republik-Ausgabe)
 MO = Der Morgen (Zeitung der LDPD/DDR)
 NN = Norddeutsche Neueste Nachrichten (Rostock, Zeitung der NDPD/DDR)
 WE = DIE WELT (Deutschlandausgabe) (B = Berliner Ausgabe)

Beispiel für den Aufbau der Informations- und Klassifikations-Datei (IK-Datei) zum Bonner Zeitungskorpus (Version 10/83)

A	B	C
BKIK 0001 ZT 01NDB69010401	ND	1B2V 1 0 114
BKIK 0002 ZT 01NDB69010401	ND	2G2Q3C13C 133
BKIK 0003FZT 01NDB69010401	ND	1A1B 23U0122
BKIK 0004 ZT 01NDB69010401	ADNND	1B10 13B 121
BKIK 0005 ZT 01NDB69010401	ND	1Q1W2G12E 122
BKIK 0006 ZT 01NDB69010401	ADNND	1M 1 A 122
BKIK 0007 ZT 01NDB69010401		3V1A 1 S 112
BKIK 0008 ZT 01NDB69010401	ND	1W1I 1 A 121
BKIK 0009FZT 01NDB69010401*Ulbricht,Walter		1B1A1D23T 141
BKIK 0011FZT 01NDB69010602		2A2G 23B 821
BKIK 0012FZT 01NDB69010602		2C3L2T1 A 821
BKIK 0013 ZT 01NDB69010602*Hempel,Wilhelm		2C3L2T12B 831
BKIK 0014 ZT 01NDB69010602 (Zeichn.Prof.*Haas)		2T 4 0 812
BKIK 0015FZT 01NDB69010602		2F2E3V12B 822
BKIK 0016 ZT 01NDB69010602*Brock,Guenther		3V1W1J12E 822
BKIK 0017 ZT 01NDB69010602*MueLLer,Erwin		3V 12E 822
BKIK 0018 ZT 01NDB69010602*Urbschat,Hans-Georg		2C3L2T12B 824
BKIK 0019 ZT 01NDB69010602*Sturzbecher,Adolf		2L2G2F12B 831
BKIK 0020 ZT 01NDB69010602	ADN	1A1B 23B 821
BKIK 0021 ZT 01NDB69010602	ND	5H5B 13B 821
BKIK 0022 ZT 01NDB69010602	ND	3B 1 A 811
BKIK 0023 ZT 01NDB69010602		3C2H 12E 832
BKIK 0024 ZT 01NDB69010602 E.Ru.		1S2G2K12E 832
BKIK 0025 ZT 01NDB69010602 BI		2L2F 23C 832
BKIK 0026 ZT 01NDB69010602		5T 13E 821
BKIK 0041 ZT 01NDB69011403		1A1R1P23T 163

In der Gruppe C sind die Felder (= Informationsklassen) 13, 14 und 16 nur bei den Bonner Jahrgängen der Phase 1 (WE und ND 54, 64 und 69) ausgefüllt; Nr. 12 und 15 (Sachgebiet und Sparte) als die wichtigsten Angaben zur Inhaltserschließung sind jedoch bei allen Zeitungsartikeln ausgefüllt.

Die Angaben in Gruppe C sind zum Teil verschlüsselt. Eine Tabelle zur Entschlüsselung siehe folgende Seiten.

In einem weiteren Informations-Satz, dem "Thema-Satz", wird außer der Artikelnummer noch das Thema, das in dem betreffenden Artikel behandelt wird, angegeben. Das geschieht in der Weise, daß zum einen die Länder genannt werden, auf die sich das im Artikel behandelte Thema bezieht, daß zum anderen das Thema in Stichworten wiedergegeben wird. Auch diese Informationen sind nur bei den Bonner Texten der Phase 1 vorhanden.

4.2. Benutzerzugriff

Ein Benutzer kann auf diese Informationen auf zweierlei Weise zugreifen:

a) Er kann sie benutzen, um den gesamten Datenbestand zu segmentieren; z.B. kann er Anzeigentexte ausschließen oder nur Agenturtexte einschließen oder das gesamte BZK 1 nach Sparten aufgliedern wollen. Dies ist durch entsprechende Selektions- und Sortierarbeiten in der IK-Datei möglich; das Ergebnis ist eine Liste von gewünschten oder ausgeschlossenen Artikelnummern, die dann als Eingabedatei für entsprechende Selektionsläufe über das Textkorpus dient.

b) Er kann aber auch, von einzelnen Textbelegen ausgehend (z.B. als Ergebnis von Suchwort-Recherchen mittels REFER) Informationen zu seinen Textbelegen wünschen; in der Regel die genaue Datierung, aber ggfs. auch die gesamten im IK-Satz gespeicherten Informationen einschließlich Sachgebiets- und Spartenzuordnung.

Hier dient die jedem Textbeleg beigegebene Artikelnummer und Jahrgangssigle als Adresse, unter der die gewünschten Informationen aus der IK-Datei abrufbar sind.

Wo kein Terminalzugang zur IK-Datei besteht, ermöglichen die unter 2.4. zusammengestellten Listen der aufgenommenen Seiten wenigstens eine genaue Datierung. In Einzelfällen kann das IdS bei der Beschaffung von Kopien der Original-Artikel helfen.

4.3. Listen zur Entschlüsselung kodierter Angaben

Zu 12) Klassifikation der Sachgebiete (vgl. Hellmann 1968, 99-103)

1 Politik

- | | |
|------------------------------------|--------------------------------------|
| 1A Allgemeine Lage | 1L Asiatische Politik (auch einzel- |
| 1B Persönlichkeiten | ner Staaten) |
| 1C Übernationale Zusammenschlüsse | 1M Arabisch-nahöstl. Politik (auch |
| und Vereinbarungen | einzelner Staaten) |
| 1D West-Ost-Politik bestimmter | 1N Deutsche Außenpolitik |
| Staaten) | 1O Deutsche West-Ost-Politik |
| 1E Westliche Außenpolitik | 1P Innenpolitik auf Bundesebene bzw. |
| 1F Östliche Außenpolitik | der DDR |
| 1G Politik der Blockfreien und | 1Q Innenpolitik auf Länderebene |
| Neutralen als Gruppe | 1R Innenpolitik auf Partei(en)ebene |
| 1H Politik einzelner blockfreier | 1S Deutsche Wirtschaftspolitik |
| und neutraler Staaten | 1T Innenpolitik bestimmter Staaten |
| 1I Militär- und Verteidigungspoli- | 1U Deutsche Kulturpolitik |
| tik (international und national) | 1V Deutsche Sozialpolitik |
| 1J Europäische Politik (auch ein- | 1W Verbände |
| zelner Staaten) | 1X Kommunalpolitik |
| 1K Afrikanische Politik (auch | 1Y |
| einzelner Staaten) | 1Z Sonstiges |

2 Wirtschaft

- | | |
|-----------------------------------|--------------------------------------|
| 2A Allgemeine Lage | 2P Handel, Innen- |
| 2B Persönlichkeiten | 2Q Betriebsorganisation, Rationali- |
| 2C Technik | sierung |
| 2D Übernationale Zusammenschlüsse | 2R Finanzen, Börse, Versicherungen |
| (außer europäischen) | 2S Verkehr, Verkehrswege |
| 2E Europäische Zusammenschlüsse | 2T Verkehrsmittel (Luft-, Land-, |
| 2F Organisationen, Verbände | Seefahrzeuge) und -technik |
| 2G Planung | 2U Stadtplanung, Raumordnung |
| 2H Arbeitskräfte, Personalstruk- | 2V Wohnungsbau, Bauwesen |
| tur | 2W Dienstleistungsindustrie und -ge- |
| 2I Industrie, Grundstoffindustrie | werbe (Fremdenverkehr, Hotel u. |
| 2J Industrie, Investitionsgüter | dergl.) |
| 2K Industrie, Konsumgüter | 2X Werbung |
| 2L Landwirtschaft, Forsten | 2Y Vergnügung, Unterhaltung, Massen- |
| 2M Seefahrt, Fischerei | medien (nur wirtschaftlich) |
| 2N Handwerk | 2Z Sonstiges |
| 2O Handel, Außen- | 2(Energiewirtschaft |
| | 2) Handel-Versorgung |

3 Soziales

- | | |
|--|---|
| 3A Allgemeines, Statistik | 3M Schul-, Erziehungs-, Bildungswesen |
| 3B Persönlichkeiten | 3N Beruf und Arbeitswelt |
| 3C Bevölkerung | 3O Haus und Garten |
| 3D Verbrechen, Vergehen (auch Beschuldigungen) | 3P Familie, (Ehe) |
| 3E Rechtswesen | 3Q Liebe (Ehe), Erotik (sog. Privat oder Intimsphäre) |
| 3F Gesundheitswesen, Medizin | 3R Hobbies, Freizeit, Reisen, Tourismus |
| 3G Verbände, Organisationen | 3S Geographisches |
| 3H Körperpflege | 3T Mode |
| 3I Versorgung (Renten usw.) | 3U |
| 3J Rassenprobleme | 3V Veranstaltungen |
| 3K Kommunikationsmittel (Presse, Rundfunk, Fernsehen, Film - nicht Artikel zum künstlerischen Rang!) | 3W Naturereignisse, Wetter |
| 3L Wissenschaft und Forschung | 3X Unglücksfälle, Todesfälle |
| | 3Y Streiks, Unruhen, Aufstände |
| | 3Z Sonstiges |

4 Sport

- | | |
|---|------------------------------|
| 4A Allgemeines | 4M Fechten |
| 4B Persönlichkeiten, Verbände | 4N Pferdesport |
| 4C | 4O Wassersport |
| 4D Großveranstaltungen mehrerer Sportarten (z.B. olymp. Spiele) | 4P Wintersport (mit Eislauf) |
| 4E Leichtathletik | 4Q Schießsport |
| 4F Schwerathletik, Boxen, Ringen | 4R Radsport |
| 4G | 4S Motor- und Flugsport |
| 4H Turnen, Gymnastik, Tanz | 4T |
| 4I | 4U Bergsteigen, Wandern |
| 4J Ballspiele, Mannschafts- | 4V |
| 4K | 4W Spiele und Rätsel |
| 4L Ballspiele, Einzel- (Tennis, Tischtennis usw.); Golf usw. | 4X |
| | 4Y |
| | 4Z Sonstiges |

5 Kulturelles

- | | |
|--|---|
| 5A Allgemeines | 5N Wissenschaft, Forschung, Universitäten |
| 5B Persönlichkeiten | 5O Massenmedien |
| 5C Museen, Galerien usw. | 5P Philosophie, Weltanschauung |
| 5D Organisationen, Gruppen | 5Q Kunstakademien |
| 5E Kulturelle Entwicklung in einzelnen Ländern | 5R Religion und Kirche |
| 5F Musik (Konzert) | 5S |
| 5G Oper, Operette, Musical | 5T Geschichte |
| 5H Theater, Ballet | 5U |
| 5I Malerei, Graphik | 5V Volks- und Brauchtum |
| 5J Bildhauerei | 5W |
| 5K Architektur | 5X |
| 5L Literatur | 5Y |
| 5M Sprache | Sonstiges, Vermischtes |

Nachbemerkung: Leere Felder sind durch spätere Zusammenlegung zweier oder mehrerer Klassen entstanden, bzw. für eventuell neu zu etablierende Klassen vorgesehen gewesen.

Dieses Sachgebietsschema ist für Zwecke der Korpus-Segmentierung vielleicht, für statistische Zwecke sicherlich schon zu fein aufgeschlüsselt. Es wurde daher eine Hyperklassifikation entwickelt, in der die einzelnen Untersachgebiete zu größeren Gruppen zusammengefaßt wurden, nämlich Politik mit 4, Wirtschaft mit 5, Soziales Leben mit 5, Sport mit 4, Kulturelles Leben mit 4 Untergruppen, also zusammen 22 Gruppen. (Zur Einteilung dieser Hyperklassifikation siehe Schaefer, Bericht II in Hellmann (Hg.) 84, Anhang S. 119-123.)

Zu 13)

Publizistische Ziele

Es können 2 Chiffren gleichzeitig angegeben werden.

1 = Unterrichtung

2 = Belehrung

3 = Beeinflussung

4 = Unterhaltung

5 = Wirtschaftswerbung

zu 14)

A = Nachricht (Unterrichtung, Wirtschaftswerbung)

B = Bericht (Unterrichtung, Belehrung, Beeinflussung, Wirtschaftswerbung)

C = Background-Bericht (Unterrichtung, Belehrung, Beeinflussung)

D = Abhandlung (Belehrung, Unterrichtung)

E = Beitrag (Belehrung, Unterrichtung, Unterhaltung, Wirtschaftswerbung)

F = Tips, Vorschlag, Anweisung (Belehrung, Unterrichtung, Unterhaltung, Wirtschaftswerbung)

G = Hauptkommentar (Beeinflussung, Belehrung, Unterrichtung)

H = Glosse (Beeinflussung, Unterhaltung)

I = "Spitzmarke" (Beeinflussung, Unterhaltung)

J = Geschichte (Unterhaltung, Belehrung) (Kurzgeschichte, Erzählung)

K = Plauderei (Unterhaltung, Beeinflussung, Belehrung)

L = "kleines Feuilleton" (Unterhaltung, Beeinflussung, Belehrung)

M = Großanzeige (Wirtschaftswerbung, Beeinflussung, Unterrichtung, Belehrung)

N = Anzeige (Wirtschaftswerbung, Beeinflussung, Unterrichtung, Belehrung)

Sonderformen

O = Text zu einem Bild, Schaubild, graph. Darstellung usw.

P = Leserbrief

Q = Interview

R = Abdruck eines literarischen Textes (auch Gedichtes)
 S = Abdruck einer amtlichen Verlautbarung (Gesetz, Vertrag, Bekanntmachung, Aufruf), Communiqué
 T = Abdruck einer Rede
 U = Abdruck eines fremden Artikels
 V = Klein- und Privatannonce
 W = Tabelle, Wetterbericht
 X = Fortsetzungsroman
 Z = Sonstiges

Für diese Sonderformen braucht kein publizistisches Ziel angegeben zu werden; die dafür vorgesehenen Spalten (64-65) werden dann mit 00 ausgefüllt.

Zu 15)

Angabe der Sparte

Politik	= 1	Lokales	= 6
Wirtschaft	= 2	Anzeigen	= 7
Sport	= 3	Vermischtes	= 8
Technik-Motor	= 4	Sonstiges	= 0
Feuilleton	= 5		

Zu 16)

Äußere Form

- 1) Länge des Artikels in Zeilen (nach Klassen gruppiert),
- 2) Breite des Artikels in Spalten.

5. MÖGLICHKEITEN DER AUSWERTUNG

5.1. Fragestellungen

Grundsätzlich erlaubt das Material von seiner Struktur her ein Herangehen unter verschiedenen Aspekten:

1. Da das Material aus Ost- und West-Zeitungen besteht: ost-west- vergleichend, und zwar sowohl generell als auch jahrgangsweise oder nach anderen Auswahlkriterien.
2. Das das Material zeitlich gestuft ist: diachronisch vergleichend, also bezogen über einen Zeitraum von 25 Jahren hin.
3. Das das Material u.a. nach Sparten und Sachgebieten klassifiziert ist

(vgl. oben 4.2): nach Sparten/Sachgebieten vergleichend, ggf. unter Einbeziehung auch des ost-west-vergleichenden und/oder des diachronisch vergleichenden Aspekts.

Außerdem ist das Material, wie oben in 3.3 beschrieben, in bestimmter Weise codiert bzw. transkribiert und kann entsprechend abgefragt werden. Man kann also z.B. die transkribierten Überschriften, Bildbeischriften, Quellenangaben aus einer Untersuchung ausschließen oder gerade gesondert hervorheben; man kann u.a. Substantive, Initialabkürzungen, großgeschriebene Adjektive abfragen sowie Satzschlüsse und andere syntaktische Binnengliederungen mithilfe codierter Satzzeichen erkennen und abfragen. Jede der in 3.3 erwähnten Codierungen erlaubt entsprechende Abfragen.

Nicht abfragbar - um auch zwei Beispiele aus der Praxis zu zitieren - sind dagegen Fragestellungen wie "Ironie im Feuilleton" oder "Manipulation im Meinungsartikel der WELT": Feuilleton und Meinungsartikel wären segmentierbar, aber "Ironie" oder "Manipulation" sind nicht maschinenexplizit, und die Anführungszeichen (z.B. vor und nach "DDR") helfen nicht weiter.

Sehr gut bearbeitbar sind dagegen in der Regel alle Fragestellungen, die sich auf Häufigkeit und Häufigkeitsrelationen von Wortformen beziehen. Gerade der Unterschied von Häufigkeiten im Zeitvergleich oder im Vergleich zwischen Sachgebieten ergibt interessante Aufschlüsse, z.B. auch über Stilerscheinungen.

Am häufigsten wird das Material zur Beschaffung von Belegen zu vorgegebenen Wörtern oder Wortbestandteilen (z.B. Morphemen) benutzt. Dazu wird eine bestimmte Zeichenkette, z.B. "demokrat", eingegeben. Der Rechner liefert dazu sämtliche Sätze, in denen Wortformen mit dieser Zeichenkette vorkommen, also auch z.B. "demokratischste, undemokratischer, demokratiefeindlich" usw. Dies Verfahren läßt sich auch bei einer Abfrage nach Morphemen verwenden, z.B. nach "-ismus", obwohl man mit Fehlbelegen wie "Kreismuster" rechnen muß. Auch Kombinationsabfragen mit (zur Zeit) 2 Gliedern sind möglich, z.B. nach "in" & "Angriff", wenn das Funktionsverbgefüge "in Angriff nehmen" bearbeitet werden soll. (Aller-

dings werden zur Zeit vom Programm auch Sätze ausgegeben, in denen die gesuchten Wörter weit entfernt stehen: "...als in diesem Augenblick der Angriff begann ..."). Die gefundenen Kontext-Belege werden zusammen mit bestimmten Nummern ausgegeben, die ein Wiederauffinden in den Texten und eine Fundstellenangabe ermöglichen. Für das Bonner Zeitungskorpus ist eine entsprechende Datierung mithilfe der in diesem Aufsatz gedruckten Artikel-Seiten-Kokordanz gewährleistet.

5.2. Ausgabeformen und -techniken

5.2.1 Ausgabe über Drucker

Die wesentlichen Ausgabeformen über Schnelldrucker sind:

- Kontext-Beleglisten als Ergebnis von Suchwort-Abfragen
- KWIC-Konkordanzen (KWIC = key word in context) oder KWIC-Auszüge, bei denen das gesuchte Wort in alphabetischer Sortierung in der Mitte von genau einer Computerzeile Kontext steht;
- Wortregister ohne Kontext, aber mit Fundstellenangaben (= Index) oder ohne letztere (= Register) und zwar in alphabetisch vorwärtsläufiger Sortierung, in alphabetisch rückläufiger Sortierung, nach Häufigkeit absteigend sortiert;
- Gesamt- oder Mischregister aus allen Jahrgängen.

Auch diese Wortregister können wieder Gegenstand von Suchwort-Abfragen sein; das Ergebnis sind entsprechende Auszüge aus den Registern.

Weitere Abfrage- und Auswertungsmöglichkeiten sollten mit der Service-Stelle im IdS, z.Hd. Frau S. Dickgießer, oder mit dem fachlich zuständigen wissenschaftlichen Mitarbeiter besprochen werden.

5.2.2 Mikrofiches

Die Texte selbst umfassen, auf Papier (Großformat) ausgedruckt, etwa 18 Bände zu je 500 bis 600 Seiten. Die Indices und Register machen etwa 30 Bände aus, die vollständigen KWIC-Konkordanzen (bei 60 Zeilen pro Seite)

152

rund 100 Bände zu je 500 Seiten. Niemand kann solche Papierberge noch speichern, geschweige denn bewegen. Auch diese Daten ständig auf Magnetplatten bereit zu halten, ist nicht immer möglich. Dafür bietet der Mikrofiche ein preiswertes und handliches Archivierungsmittel. In 48-facher Verkleinerung speichert ein postkartengroßer Mikrofiche 270 Computerseiten; 100 Bände Großformat lassen sich auf weniger als 200 Mikrofiches unterbringen.

Das gesamte Bonner Zeitungskorpus Teil 1 mit allen Registern und Konkordanzen paßt bequem in eine Aktentasche. Ein Lesegerät kostet weniger als ein Bücherregal. Im IdS und in einigen großen Bibliotheken sind reader-printer vorhanden, mit denen man die (ausgewählten bzw. benötigten) Daten vom Mikrofiche wieder rückvergrößern und auf Papier kopieren kann. Bei einer Reihe von Fragestellungen bzw. Arbeitsgängen macht die Verwendung von Mikrofiches und reader-printer die Benutzung des Rechners überhaupt überflüssig.

Das BZK 1 erscheint auf Mikrofiche unter folgendem Titel:

Das Bonner Zeitungskorpus Teil 1. Am Institut für deutsche Sprache erstellt nach der Konzeption und unter der Leitung von Manfred W. Hellmann. Teil 1: DIE WELT und NEUES DEUTSCHLAND. Texte, Register, Konkordanzen, Gesamtregister. (= Brekle u.a. (Hg.), Regensburger Mikrofiche Materialien (RMM) Nr. 007/1), Regensburg 1985. (MCS Verlag Nürnberg).

5.2.3 Bildschirm-Auswertung

Im IdS selbst ist ein dialogorientiertes, sehr benutzerfreundliches Abfragesystem namens REFER entwickelt worden, mit dessen Hilfe der Benutzer in verschiedener Hinsicht seine Abfragen selbst steuern, Belege kommentieren, abspeichern oder ausgeben lassen kann. Eine detaillierte Beschreibung ist bei der IdS-Service-Stelle erhältlich, sie ersetzt allerdings das Ausprobieren nicht. REFER enthält auch die oben erwähnte Möglichkeit der Kombinationsabfrage. Vor allem ist es sehr schnell, d.h. die Antwortzeiten des Systems am Bildschirm sind erfreulich kurz. REFER wird in gewissen Zeitabständen weiterentwickelt.

5.3. Bearbeitung von Anwenderproblemen

Es kommt vor, daß Anwenderprobleme mit den vorhandenen Bereitstellungsformen und Auswertungsprogrammen einschließlich REFER nicht zu lösen sind. Je nachdem ob es sich eher um ein linguistisches oder um ein technisches Problem handelt, sollte der zuständige wissenschaftliche Mitarbeiter im IdS oder aber die Service-Stelle oder beide mit dem Problem vertraut gemacht werden. Manchmal findet sich noch ein Lösungsweg. Manchmal können wir auch ein Anwenderproblem zu unserem eigenen machen und finden gemeinsam eine neue Lösung. Es lohnt sich also nachzufragen (vgl. auch M. Kolvenbach in IdS LDV-Info 1, 1981, S. 1-4).

6. AUSBLICK

6.1 Fortschreibung

Das Bonner Zeitungskorpus ist auf Fortschreibung in bestimmten Intervallen hin angelegt. Nur so läßt sich laufend Aktualität gewährleisten, und nur durch Aktualität bleibt das Korpus auch in seinen früheren Querschnitten interessant. Der Jahrgang 1979 wurde daher schon kurz nach der Übersiedlung von Bonn nach Mannheim ausgezählt, die Auswahl fotokopiert und zur Texterfassung vorbereitet. Der Jahrgang 1984 steht jetzt an. Das IdS sucht nach einer technischen Lösung, die die immensen Kosten einer manuellen Texterfassung auf ein erträgliches Maß reduziert.

6.2. Erweiterungen

Das Bonner Zeitungskorpus enthält mit der WELT und dem ND gewiß zwei für den jeweiligen Staat durchaus typische Zeitungen, trotzdem würde man sich, zumindest für die westdeutsche Seite, Ergänzungen wünschen. Sie wurden noch in Bonn durch die Aufnahme folgender Jahrgangsauswahlen vorgenommen:

Bonner General-Anzeiger 1964

Bonner General-Anzeiger 1974

Frankfurter Rundschau 1974

Rostocker Neueste Nachrichten (NDPD/DDR) 1964

Rostocker Neueste Nachrichten (NDPD/DDR) 1974

Der Morgen (LDPD/DDR) 1974.

Diese Jahrgangsauswahlen bilden den Teil 2 des BZK. Sie sind nach dem gleichen Prinzip aufgenommen wie Teil 1. Sie werden allerdings nicht mehr den Mannheimer Codierungskonventionen angepaßt und bleiben im wesentlichen unkorrigiert. Sie sind mittels allgemeiner Suchwortprogramme auswertbar, jedoch nicht mittels REFER.

An eine Fortschreibung auch dieser Erweiterungstexte wird zur Zeit nicht gedacht.

Statt ihrer entsteht zur Zeit - im Rahmen eines großen lexikologischen Vorhabens - eine umfangreiche Sammlung maschinenlesbarer Texte des "Mannheimer Morgen" Jg. 1985. Die Texte sind jedoch vorläufig nicht verfügbar.

Literatur

Bayer, Klaus/Kurbel, Karl: Maschinelle Textverarbeitung im Rechenzentrum des Instituts für deutsche Sprache. In: Bayer/Kurbel/Epp: Maschinelle Textverarbeitung im Rechenzentrum des IdS (= Forschungsberichte des IdS Nr. 14). Tübingen 1974, S. 1-66.

Hellmann, Manfred W.: Zur Dokumentation und maschinellen Bearbeitung von Zeitungstexten in der Außenstelle Bonn des Instituts für deutsche Sprache (= Forschungsberichte des Instituts für deutsche Sprache Nr. 2). Mannheim 1968, S. 39-125.

Hellmann, Manfred W.: Untersuchungen an östlichen und westlichen Zeitungstexten. Zu einigen Arbeiten der Außenstelle Bonn des Instituts für deutsche Sprache. In: H. Schanze (Hg.), Literatur und Datenverarbeitung. Tübingen 1972, S. 66-70.

Hellmann, Manfred W.: Bericht über die Arbeit der Forschungsstelle Bonn des Instituts für deutsche Sprache. In: Hellmann, Manfred W. (Hg.): Zum öffentlichen Sprachgebrauch in der Bundesrepublik Deutschland und in der DDR - Methoden und Probleme seiner Erforschung. (= Reihe Sprache der Gegenwart Bd. 18). Düsseldorf 1973, S. 15-34.

Hellmann, Manfred W.: Sprache zwischen Ost und West - Überlegungen zur Wortschatzdifferenzierung zwischen BRD und DDR und ihren Folgen. In: Kühlwein, W./Radde, G. (Hg.): Sprache und Kultur: Studien zur Diglossie, Gastarbeiterproblematik und kulturellen Integration (= Tübinger Beiträge zur Linguistik Bd. 107). Tübingen 1978, S. 15-54 (mit einem Anhang Wortläufigkeitstabellen, S. 50 ff).

Hellmann, Manfred W. (Hg.): Ost-West-Wortschatzvergleiche. Maschinell gestützte Untersuchungen zum Vokabular von Zeitungstexten aus der BRD und der DDR. (= Forschungsberichte des IdS Bd. 48). Tübingen 1984.

Hellmann, Manfred W.: Das Projekt "Ost-West-Wortschatzvergleiche" - Neue Wege zur Untersuchung der sprachlichen Ost-West-Differenzen. In: M. W. Hellmann (Hg.), 1984, S. 11-73.

Hellmann, Manfred W.: Das Bonner Zeitungskorpus Teil 1. In: IdS LDV-Info 4. Mannheim 1984, S. 38-47.

Institut für deutsche Sprache: Das Bonner Zeitungskorpus Teil 1. Am Institut für deutsche Sprache erstellt nach der Konzeption und unter der Leitung von Manfred W. Hellmann. Teil 1: DIE WELT und NEUES DEUTSCHLAND. (= Brekle u.a. (Hg.), Regensburger Microfiche Materialien (RMM) Nr. 007/1). Regensburg 1985 (MCS Verlag Nürnberg). (Microfiche-Veröffentlichung der Texte, Register und Konkordanzen).

Mückenmüller, Peter: Die maschinelle Aufbereitung des Bonner Zeitungskorpus Teil 1: In IdS LDV-Info 4. Mannheim 1984, S. 48-56 (mit Muster-Auswertungen).

Institut für deutsche Sprache, Abt. Zentrale Wiss. Dienste: IdS LDV-Info 1: Nachrichten aus der Arbeitsstelle Linguistische Datenverarbeitung. (masch.), Mannheim 1981
IdS LDV-Info 2, Mannheim 1982
IdS LDV-Info 3, Mannheim 1983
IdS LDV-Info 4, Mannheim 1984
(jeweils Eigenverlag des IdS).

Schaeder, Burkhard: Eine Dokumentation maschinenlesbarer Textkorpora der deutschen Gegenwartssprache. In: Henne u.a. (Hg.) 1978, S. 233-254.

Schaeder, Burkhard: Maschinelle Dokumentation und Lexikographie. Ausführungen zum DFG-Projekt Ost-West-Wortschatzvergleiche. In: Krallmann (Hg.) 1978, S. 110-164.

Schaeder, Burkhard: Zur Methodik der Auswertung von Textkorpora für die Zwecke der Lexikographie. In: Bergenholtz/Schaeder (Hg.), Empiri-

sche Textwissenschaft. Aufbau und Auswertung von Textkorpora. Königstein/Ts 1979, S. 220-267.

Schaeder, Burkhard: Maschinenlesbare Text-Corpora des Deutschen und des Englischen. Eine Dokumentation. In: Bergenholtz/Schaeder (Hg.) 1979, S. 325-336 (Anhang).

Schaeder, Burkhard: Lexikographie als Praxis und Theorie (= Reihe Germanistische Linguistik Bd. 34). Tübingen 1981.

Schaeder, Burkhard: Das Bonner Zeitungskorpus. Eine maschinelle Dokumentation von Tageszeitungen aus der BRD und der DDR. In: M.W. Hellmann (Hg.). 1984, S. 74-123 (Bericht Nr. II).